

THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTE UNIVERSITE
GRENOBLE ALPES**

Spécialité : **Génie Electrique**

Arrêté ministériel : 7 août 2006

Présentée par

Audrey LE MOUNIER

Thèse dirigée par **Benoît DELINCHANT** et
codirigée par **Stéphane PLOIX**

préparée au sein du **Laboratoire de Génie Electrique de
Grenoble (G2Elab)** et du laboratoire des Sciences pour la
conception, l'Optimisation et la Production (G-SCOP)
dans l'École Doctorale d'Electronique, Electrotechnique,
Automatique et Traitement de Signal (EEATS)

Méta-optimisation pour la calibration automatique de modèles énergétiques bâtiment pour le pilotage anticipatif

Thèse soutenue publiquement le « **29 Juin 2016** »,
devant le jury composé de :

Madame Florence OSSART,
Professeur à l'UPMC, Sorbonne université, Rapporteur

Monsieur Christian INARD,
Professeur à l'université de la Rochelle, Rapporteur

Monsieur Alexandre NASSIOPOULOS
Chercheur à l'IFSTTAR, Examineur

Monsieur Benoît DELINCHANT,
Maître de conférences à l'UGA, Directeur de thèse

Monsieur Stéphane PLOIX,
Professeur à Grenoble INP, Co-directeur de thèse



à Papy Vichy et Papy Rolleboise,

Table des matières

| | | |
|-----------|---|-----------|
| | Remerciements | 10 |
| | Résumé | 13 |
| | Introduction générale | 16 |
| I | Contexte, problème et état de l’art | 17 |
| | I.1 Le contexte énergétique : bilan et enjeux | 18 |
| | I.2 La gestion anticipative | 20 |
| | I.2.1 Les approches sans modèle et leurs limites | 20 |
| | I.2.2 Les approches à base de modèles | 21 |
| | I.3 Obtention des modèles pour la gestion | 21 |
| | I.3.1 Identification des systèmes | 21 |
| | I.3.2 Modélisation thermique des bâtiments | 24 |
| | I.4 Problématiques de calibration des modèles | 27 |
| | I.4.1 Les sources d’incertitudes agissant sur un bâtiment | 27 |
| | I.4.2 Problématique de la thèse : Robustesse en identification et recalibrage des modèles | 28 |
| II | Mise en évidence sur un cas d’étude des limites des approches à base de modèle d’observation | 31 |
| | II.1 Description du cas d’étude : les salles PREDIS MHI | 33 |
| | II.1.1 Description générale | 33 |
| | II.1.2 Mesures et simulation | 34 |
| | II.2 Application d’approches d’identification type séries temporelles | 42 |
| | II.2.1 Les modèles universels | 42 |
| | II.2.2 Résultats d’identification | 43 |
| | II.3 Instabilité des modèles | 51 |
| | II.4 Besoin de modèles non-linéaires | 52 |

| | | |
|------------|--|-----------|
| III | Première approche de l'estimation paramétrique à base de modèle physique | 53 |
| III.1 | Avantages des approches à base de modèle physique | 55 |
| III.1.1 | Intègre la connaissance disponible sur le système | 55 |
| III.1.2 | Extrapolabilité et capacité à représenter des situations non représentées dans le jeu de données | 55 |
| III.1.3 | Capacité du modèle à être validé par la physique | 56 |
| III.2 | Description du modèle physique de PREDIS MHI | 56 |
| III.2.1 | Confort des occupants | 57 |
| III.2.2 | Modélisation thermique de PREDIS MHI | 60 |
| III.2.3 | Les entrées du modèle : températures et puissances injectées | 64 |
| III.2.4 | Les paramètres du modèle | 69 |
| III.3 | Estimation paramétrique en utilisant un modèle physique selon une approche naïve | 71 |
| III.3.1 | Simulation du modèle avec les valeurs initiales des paramètres | 71 |
| III.3.2 | Première estimation des paramètres du modèle | 72 |
| III.4 | Problématiques liées à l'estimation paramétrique en utilisant un modèle physique | 73 |
| III.4.1 | Approche non-linéaire | 73 |
| III.4.2 | Besoins de valeurs de paramètres initiales | 75 |
| III.4.3 | Identifiabilité | 75 |
| III.4.4 | Obtention d'un modèle physique adapté | 78 |
| IV | Analyse des problématiques liées à l'estimation paramétrique à base de mo- dèle physique | 79 |
| IV.1 | Estimation des paramètres | 81 |
| IV.1.1 | Identifiabilité des paramètres | 82 |
| IV.1.2 | Analyse de sensibilité | 83 |
| IV.1.3 | Analyse de la dispersion paramétrique et de l'arrivée en butée des va- leurs identifiées | 88 |
| IV.2 | Étude de la richesse spectrale des signaux d'entrée | 89 |
| IV.2.1 | Sollicitation de la pompe à chaleur de l'espace bureau à l'aide d'une séquence binaire pseudo-aléatoire | 91 |
| IV.2.2 | Longueur des données nécessaires à l'identification | 95 |
| IV.2.3 | Pertinence du choix des bornes sur les résultats d'estimation paramétrique | 98 |

| | |
|---|------------|
| IV.3 Incertitudes sur les phénomènes | 99 |
| IV.3.1 Influence de l'état d'ouverture des portes sur l'identification de la température intérieure de PREDIS MHI | 100 |
| IV.3.2 Influence des scénarios de ventilation sur l'identification des salles PREDIS MHI | 101 |
| IV.4 Besoin de "méta-optimisation" | 102 |
| IV.4.1 Amélioration de l'identifiabilité paramétrique | 102 |
| IV.4.2 Etude des phénomènes mal maîtrisés | 103 |
| V Analyse de la dispersion paramétrique afin d'obtenir un modèle réduit permettant un recalibrage robuste | 105 |
| V.1 Description du principe | 107 |
| V.1.1 Algorithme | 107 |
| V.1.2 Scénario de mise en place | 108 |
| V.1.3 Besoins et pré-requis techniques | 110 |
| V.2 Mise en place de la méthode sur des jeux de données PREDIS MHI | 111 |
| V.2.1 Test de la méthode sur des jeux de données obtenus par simulation COMFIE Pléiade | 111 |
| V.2.2 Test de la méthode sur des jeux de données expérimentaux | 116 |
| V.3 Validation de la méthode avec un autre modèle de bâtiment | 126 |
| V.3.1 Description du modèle | 126 |
| V.3.2 Description du modèle physique : utilisation d'un réseau de résistances et capacités thermiques | 126 |
| V.3.3 Mise en place de la méthode d'analyse de la dispersion | 129 |
| V.4 Simulation de la méthode en scénario de gestion | 131 |
| V.5 Conclusions sur l'utilisation de la méthode de réduction par analyse de la dispersion paramétrique | 132 |
| V.5.1 Avantages et inconvénients de cette méthode | 132 |
| V.5.2 Conclusion sur la méthode et généralisation | 133 |
| VI Pilotage optimal de l'estimation paramétrique par méta-optimisation (systématisation) | 135 |
| VI.1 Présentation de la méta-optimisation | 136 |
| VI.1.1 Intérêt de l'utilisation de la méta-optimisation | 136 |

| | | |
|----------|--|------------|
| VI.1.2 | Description du principe | 137 |
| VI.1.3 | Algorithme | 139 |
| VI.1.4 | Pré-requis | 140 |
| VI.2 | Mise en place de la méthode sur des jeux de données PREDIS MHI | 141 |
| VI.2.1 | Simulation sur un an de jeux de données simulés issus du logiciel COMFIE | 141 |
| VI.2.2 | Validation de la méta-optimisation sur les jeux de données expérimentaux des salles PREDIS MHI | 144 |
| VI.3 | Validation de la méthode avec un autre modèle de bâtiment | 149 |
| VI.4 | Simulation de la méthode en scénario de gestion | 149 |
| VI.5 | Conclusion | 151 |
| VI.6 | Couplages des deux méthodes | 152 |
| VI.6.1 | Principe | 152 |
| | Conclusion générale | 158 |
| A | Calcul des coefficients d'interface | 159 |
| B | Calculs initiaux des résistances et capacités thermiques pour PREDIS MHI | 161 |
| B.1 | Détails des calculs des résistances et capacité thermiques du modèle PREDIS MHI | 162 |
| | Bibliographie | 165 |
| | Table des figures | 171 |

Remerciements

En premier lieu, mes remerciements s'adressent bien évidemment aux membres du jury qui ont accepté d'évaluer mes travaux de thèse.

Merci à Madame Florence OSSART, Professeur à l'UPMC, Sorbonne université et à Monsieur Christian INARD, Professeur à l'université de la Rochelle d'avoir rapporté mes travaux. Merci pour vos retours et vos questions constructives qui ont contribué à la version finale de ce manuscrit.

Merci également à Monsieur Alexandre NASSIOPOULOS, chercheur à l'IFSTTAR, d'avoir accepté d'examiner mes travaux, c'était un grand plaisir de vous compter parmi les membres du jury.

Merci à mes deux papas qui m'ont encadrée dans la réalisation de cette thèse. Stéphane, merci pour ta créativité scientifique, tes idées ingénieuses et ton immense savoir. Benoît, je te remercie pour ton écoute, ta patience, ta tolérance et ta rigueur scientifique. Grâce à vous, j'ai pu réaliser cette thèse dans des conditions optimales. Vous avez très bien joué votre rôle de parents, jamais d'accord, mais toujours dans la recherche de mon bien-être. Merci d'être aussi passionnés !

Les conséquences de cette garde alternée vont me faire remercier les deux laboratoires entre lesquels j'ai partagé mon temps pendant ces trois années.

Merci au G-SCOP ainsi qu'à tous ses membres. Je voudrais remercier ici quelques personnes qui m'ont marqué par leur rencontre. Lucie et Ingwild, dommage qu'on se soit rencontrées si tard, mais je suis contente de continuer à vous côtoyer de temps à autre. Lisa, merci pour ta joie de vivre et pour nos échanges, tant sur le plan scientifique qu'humain ! Ma petite Khadija, tu es et resteras une personne exceptionnelle avec qui je suis très heureuse d'avoir partagé mon bureau. Tu es la personne la plus douce, la plus ouverte et la plus aimable que je connaisse, ne change surtout pas. Et bon courage pour la fin !

Merci également à l'ensemble du G2Elab. Au club théâtre pour les bons moments passés ensemble : Laurence, Sylvie, Sylvianne, Damien, Fabien, Diandra, Maël, Pich, Laura, Maximin. A tous les copains du début : Raphaël, Archie, Victor, Manue, Aurélien, Gasp', Te, Justine, Jean-Louis, Clémentine, Guillaume, et à tous ceux de la fin : Arnaud, Quentin, Melissa (super coloc!), Thibaut, Léo. Merci aussi aux deux pièces rapportées qui « égayaient » nos pauses midis : Corentin et Stéphane ! Un merci tout particulier à Rozenn et Vincent ! Merci pour les discussions, les rigolades et les sorties nocturnes. Ma tête s'en souvient encore !

Lorsqu'on s'attèle à une thèse, il est important d'avoir un bon cadre de travail, ce qui fût le cas, mais il est également primordiale d'avoir l'occasion de se changer les idées, ce qui fût largement fait grâce à :

La danse ! Merci aux fiiiilles : Mélanie, Solenn et Céline, pour les bons moments (chorégraphiques et pas que) partagés. Merci à tous les membres de la team Savoy Cup : Zouky, Aurore, Vivien, Aude, Pom, Mary, Matthia, Céline, Solenn pour ce projet complètement fou et motivant. Merci aussi à la troupe et à mes élèves ! Merci aux swingueurs grenoblois : Léa (merci pour les discussions intéressantes et la piscine de tes parents;)), Jean-Yves et Martine (mes parents de substitutions), les M3JF : Matthieu, Matthia, Marie, Julia et François

(merci spécial pour les blues, les balboas et l'oreille attentive), Julien (merci pour les cours et les danses), Pierro (partenaire!! ou presque :p), Ludo, Aurélie, Aurélien, Xavi, Guillaume, et merci à tous les autres membres de Grenoble Swing! Mais aussi aux swingueurs non grenoblois : Jeanne, Mélanie, Aurélie, Éric, Maëlle, Hugo, Peter, Chloé, Malika, Simon : c'est toujours un plaisir de vous croiser en festival!!

Mais aussi l'acroyoga! Merci à Max pour ta... solidité? :) A tous les sens du terme! Arnaud pour nos débuts chaotiques, nos discussions interminables, ta douceur et ta tendresse. Merci Estelle, pour ta voix et ton insouciance. Merci Yves, de me permettre d'ouvrir un peu plus mon esprit chaque jour, et pour les moments agréables en vol et en base :). Merci Patrick, pour ton humeur inébranlable : festive, explosive et entière! Julien, pour tes yeux bleus qui parlent sans que tu n'ouvres la bouche, et ton sourire!

Merci aux copains de l'école. Tous les copains complètement crazy : Jean, Elie, Justine (et Tom et Gaston), Barbu, Maëlle, Trou, Laurianne, et les autres. Un merci tout particulier pour Jérémy, amoureux à temps partiel mais ami à temps complet, merci de m'aimer exactement pour ce que je suis, et de m'aider à corriger les erreurs, de n'importe quelle nature;)! Merci à Antho et merci à la famille : Yann et Yvan!

Enfin, un grand merci à ma famille, ceux qui sont encore là et les autres. Papa et Maman, merci pour tout votre soutien bienveillant, matériel et immatériel!

Résumé

Face aux enjeux climatiques actuels, le secteur bâtiment est encouragé à réduire sa consommation énergétique tout en préservant le confort des occupants. C'est dans ce contexte que s'inscrit le projet ANR PRECISION qui vise au développement d'outils et de méthodes pour la gestion énergétique optimisée des bâtiments qui nécessitent l'utilisation de modèles thermiques dynamiques. Les travaux de thèse, effectués entre le G2Elab et le G-SCOP, se sont focalisés sur les problématiques liées à l'estimation paramétrique de ces modèles. En effet, les incertitudes liées aux phénomènes mal maîtrisés et la nature des modèles rendent le calibrage des paramètres des modèles délicat. Cette procédure complexe n'est à ce jour pas systématisable : les modèles auto-regressifs ont une faible capacité d'extrapolation car leur structure est inadaptée, tandis que les modèles physiques sont non-linéaires par rapport à de nombreux paramètres : les estimations conduisent à des optimums locaux fortement dépendant de l'initialisation. Pour lever ce verrou, plusieurs approches ont été explorées à partir de modèles physiques adaptés pour lesquels des études sur l'identifiabilité ont été menées sur une plateforme expérimentale : PREDIS MHI. Différentes stratégies d'optimisation sont alors proposées visant à déterminer les paramètres qui peuvent être recalés. La première approche repose sur une analyse a priori de la dispersion paramétrique, la seconde repose sur une procédure de méta-optimisation qui détermine dynamiquement, au fur et à mesure d'une séquence d'optimisations, les paramètres à recalculer. Les résultats sont analysés et comparés à diverses approches (modèles universels, identification « naïve » de tous les paramètres d'un modèle physique, algorithme génétique, ...) à travers différents cas d'application.

In order to tackle the actual climate issues, the building field is encouraged to reduce his energetic consumption without changing the occupant's comfort. In this context, the aim of the ANR PRECISION project is to develop tools and methods for energetic management of the buildings which needs the use of dynamical thermal models. The PHD works, realised between the G2Elab and the G-SCOP, was focused on models parametric estimation issues. Indeed, uncertainties due to unknown phenomena and the nature of models lead to difficulties for the calibration of the models. Nowadays, this complex procedure is still not automatable : auto-regressive models have a low capacity to extrapolate because of their inadequate structure, whereas the physical models are non-linear regarding many parameters : estimations lead towards local optimums which highly depend on the initial point. In order to eliminate these constraints, several approaches have been explored with physical models adapted for which identifiability studies have been reached on an experimental platform : PREDIS MHI. Different optimisation strategies will be proposed in order to determine the parameters which can be estimated. The first approach uses an analyse a priori of the parametric dispersion, the second one use a meta optimisation which dynamically determined as the optimisation sequence, the parameters which can be readjusted. The results are analysed and compared to several approaches (universal models, "simple" identification of all the parameters of a physical model, genetic algorithm ...) in different application cases.

Introduction générale

Du fait de sa part non négligeable dans le bilan énergétique globale, le secteur bâtiment est devenu le cheval de bataille de la recherche. En effet, afin de répondre aux enjeux climatiques actuels, il s'agit de réduire la consommation énergétique ainsi que les émissions de gaz à effet de serre. Pour cela, des normes et directives ont été mise en place afin d'atteindre ces objectifs de réduction. Le protocole Kyoto, notamment, est un protocole international qui a visé à réduire les émissions de gaz à effet de serre entre 1990 et 2012. Depuis janvier 2013, en France, tous les bâtiments neufs doivent répondre à la RT2012 qui est une réglementation qui cadre la thermique des bâtiments en terme d'efficacité et de consommation énergétique.

Dans mes travaux de thèse, nous nous sommes intéressés plus particulièrement à la gestion anticipative. Il s'agit d'anticiper la température intérieure afin de la réguler de manière à maximiser le confort de l'occupant tout en limitant la consommation énergétique. Diverses approches peuvent être employées dans ce but. En effet, la régulation thermique peut être réalisée sans modèle de bâtiment, à l'aide d'un thermostat physique. Cependant, ce type d'approche ne permet pas une régulation multi-objectifs et se focalise uniquement sur la consigne de température au détriment de l'économie d'énergie. La gestion anticipative vise à prédire le comportement thermique du bâtiment afin de gérer de manière intelligente les apports internes gratuits et disponibles (flux solaire, flux de chaleur émis par les appareils électriques ou par les personnes) dans une pièce que l'on souhaite par exemple chauffer. Cette anticipation permet de mettre en place des scénarios des appareils de chauffage et de ventilation de manière optimale d'un double point de vue : celui du confort de l'occupant et de l'énergie consommée. C'est pourquoi les régulateurs actuels utilisent de plus en plus des modèles thermiques dynamiques. Mais actuellement, les écarts entre mesures et simulations sont encore trop importants.

C'est dans ce contexte que de nombreux projets de recherche ont émergés, notamment le projet PRECCISION qui vise le développement d'outils et de méthodes pour le pilotage des bâtiments réhabilités. Mes travaux de thèse se sont focalisés sur l'estimation paramétrique des modèles thermiques de bâtiments. Il s'agit de fournir un modèle de bâtiment qui réponde aux enjeux de la gestion anticipative, en particulier qui permette un recalibrage rapide (modèle léger) et robuste (bonne précision en prédiction). En termes de modélisation, il existe plusieurs types de modèles dont on peut distinguer deux grandes catégories : ceux qui sont inspirés de la physique et les autres, comme les modèles universels. Ces derniers, tel que les modèles ARX ou encore ARMAX donnent de très bons résultats en identification. Les modèles physiques quant à eux, possèdent une grande capacité d'adaptation lorsque les sollicitations d'un jeu de données à valider ne sont pas exactement similaires à celles du jeu de données ayant servis à l'identification des paramètres (scénarios de ventilation différent par exemple). C'est ce genre de modèles que nous avons donc utilisés dans le cadre de la gestion anticipative.

Mes travaux se sont appuyés sur les salles PREDIS MHI, un bureau et une salle de classe situés dans les anciens bâtiments du G2Elab (laboratoire de génie électrique) sur le campus de Grenoble. A l'aide de capteurs, des mesures de températures, de consommations électriques et de débit de ventilation, entre autres, ont été enregistrés. Ces deux salles ont été modélisées par des schémas électriques équivalents composés de résistances et de capacités thermiques. Les variables obtenues par capteurs ont servis d'entrée pour ces modèles. Les

problématiques liées à l'estimation paramétrique ont été appréhendées par des analyses de sensibilité. Il s'agit d'étudier la complexité du modèle afin de la corrélérer à la richesse des signaux d'entrée disponibles. Un mauvais choix dans la complexité du modèle peut mener à un manque d'identifiabilité des paramètres. En effet, ceux-ci peuvent donner des résultats dispersés à l'issue d'une série d'identification.

Ainsi, dans un premier temps, nous avons tenté d'améliorer la robustesse du procédé d'identification en agissant directement sur le modèle. En effet, une calibration classique présente des limites et offre des résultats dispersés (valeurs des paramètres et de l'erreur), ainsi que des résultats arrivant aux bornes du domaine de validité. Dans un premier temps, à travers une analyse de la dispersion paramétrique, le modèle a été réduit et les paramètres jugés non identifiables (trop dispersés) ont été fixés à leurs valeurs nominales pendant la phase d'estimation. Cette méthode cherche à résoudre les problèmes d'identification du modèle. Cette méthode a été testée et validée sur le modèle des salles PREDIS MHI puis sur un autre modèle de bâtiment. Dans un second temps, nous nous sommes focalisés sur la cohérence des valeurs estimées des paramètres. Pour cela, une méthode d'estimation pilotée afin d'éviter les zones physiquement incohérentes a été testée. Il s'agit d'une méta-optimisation qui contraint les valeurs des paramètres estimés à des valeurs qui conservent la cohérences physiques du modèle. La validité physique peut conduire à une plus grande capacité d'extrapolation du modèle à d'autres situations, car les zones de l'espace de recherche trop éloignées de la réalité pourrait donner ponctuellement de bons résultats d'identification mais être mauvais en prédiction. Cette méthode a également été testée et validée sur le modèle des salles PREDIS MHI ainsi que sur un autre modèle de bâtiment.

CHAPITRE I

Contexte, problème et état de l'art

Table des matières

| | | |
|-------|---|----|
| I.1 | Le contexte énergétique : bilan et enjeux | 18 |
| I.2 | La gestion anticipative | 20 |
| I.2.1 | Les approches sans modèle et leurs limites | 20 |
| I.2.2 | Les approches à base de modèles | 21 |
| I.3 | Obtention des modèles pour la gestion | 21 |
| I.3.1 | Identification des systèmes | 21 |
| I.3.2 | Modélisation thermique des bâtiments | 24 |
| I.4 | Problématiques de calibration des modèles | 27 |
| I.4.1 | Les sources d'incertitudes agissant sur un bâtiment | 27 |
| I.4.2 | Problématique de la thèse : Robustesse en identification et recalibrage des modèles | 28 |

I.1 Le contexte énergétique : bilan et enjeux

D'ici 2020, la France et l'Europe se sont engagées à améliorer fortement l'efficacité énergétique, développer la part des énergies renouvelables dans la consommation d'énergie et réduire leurs émissions de gaz à effet de serre et de polluants atmosphériques. Or, le secteur bâtiment représente 44 % du bilan énergétique français [2]. C'est le secteur le plus consommateur d'énergie. Il représente également 20 % des émissions nationales de gaz à effet de serre. Cette part de consommation du bâtiment ne cesse d'augmenter, par exemple, les logements et les bureaux ont vu leur consommation augmenter de plus de 30 % ces 30 dernières années. Cette augmentation est liée :

- à la croissance permanente de la taille du parc considéré ;
- à l'augmentation de la surface moyenne occupée ;
- à l'accroissement de l'exigence du confort ;
- et à l'émergence de plusieurs nouveaux usages électriques (ordinateurs, appareils électroniques, climatiseur...)

Dans le cas particulier du secteur résidentiel, le chauffage électrique représente la plus grande part de la consommation (supérieure à 60 % de la consommation totale).

Ces dernières années, une prise de conscience environnementale globale a orienté la recherche afin de mieux concevoir les bâtiments. C'est pourquoi la conception à haute performance énergétique a motivé plusieurs projets de recherche. En effet, les ressources naturelles non renouvelables s'épuisent avec le temps, entraînant une hausse importante du coût de l'énergie.

A l'échelle de l'Europe, plusieurs directives se sont intéressées à ce sujet. La directive 2002/91/CE sur la performance énergétique du bâtiment explique les engagements de l'Union Européennes dans le cadre du protocole de Kyoto, un traité international visant à la réduction des émissions de gaz à effet de serre entre 1990 et 2012 signé le 11 décembre 1993. Elle a conduit les états membres à adopter de nouvelles règles en matière d'isolation thermique et de ventilation. Deux autres directives l'ont remplacé depuis : les directives 2010/30/UE et 2010/31/UE qui concernaient la consommation en énergie et la performance énergétique des bâtiments.

En France, depuis janvier 2013, tous les bâtiments neufs doivent répondre à la RT2012 dont les exigences imposées sont de trois types :

- **L'efficacité énergétique du bâtiment** : limitation simultanée du besoin en énergie pour les composantes liées à la conception du bâtiment, imposant ainsi son optimisation indépendamment des systèmes énergétiques mis en œuvre ;
- **La consommation énergétique du bâtiment** : la consommation maximale ne doit pas dépasser un coefficient "Cepmax" modulé selon la localisation géographique, l'altitude, le type d'usage du bâtiment et la surface moyenne des logements. Cette exigence impose également le recours à des équipements énergétique performants, à haut rendement.

- **Le confort d'été dans les bâtiments non climatisés** : La RT 2012 définit des catégories de bâtiments dans lesquels il est possible d'assurer un bon niveau de confort en été sans avoir à recourir à un système actif de refroidissement. Pour ces bâtiments, la réglementation impose que la température la plus chaude atteinte dans les locaux, au cours d'une séquence de 5 jours très chauds d'été n'excède pas un seuil égal au TIC (Température Intérieure Conventionnelle).

Ainsi la RT2012 limite la consommation et permet d'éviter les projets techniquement aberrants mais le réalisme de la simulation liée à cette réglementation reste limité. En effet, les scénarios d'occupation sont fixés et indépendants des conditions de confort et les stratégies de gestion énergétique liées à l'occupation ne peuvent pas être prises en compte.

De nombreuses études ([34]) ont montré des écarts importants entre la simulation et la consommation réelle des bâtiments. Ces écarts sont dus à de nombreuses incertitudes qui interviennent pendant différentes phases de construction du bâtiment (conception, construction et exploitation) mais également aux bonnes ou aux mauvaises pratiques qui sont adoptées dans un bâtiment. Ainsi, une bonne pratique peut réduire la consommation d'énergie de 15 à 29% tandis qu'une mauvaise pratique peut l'augmenter de 49 à 79% par rapport aux prévisions [37].

De plus, plus la conception vise la basse consommation et plus les interactions et les effets de couplage entre le bâtiment, son environnement et ses usages deviennent importants. Les performances réelles dépendent fortement de la manière dont le bâtiment est utilisé et géré alors il devient difficile de définir intuitivement un mode de fonctionnement optimal. Pour assurer à tout instant le meilleur confort possible tout en évitant de dégrader les performances du bâtiment, celui-ci doit s'adapter en continu aux usages et conditions environnementales changeantes tout en anticipant l'évolution de son état futur et de différents facteurs tels que les conditions météorologiques, les tarifs ou le contenu CO_2 de l'électricité, etc.

C'est dans ce contexte que s'inscrit le projet PRECISION qui vise le développement d'outils et de méthodes pour le pilotage des bâtiments réhabilités afin de répondre à ce besoin. Ces outils accompagneront un bâtiment qui vient de subir une réhabilitation en donnant au gestionnaire, à l'occupant ou aux autres acteurs impliqués des indicateurs et les moyens numériques pour analyser le comportement du bâtiment et en contrôler le fonctionnement, soit en agissant directement sur les systèmes à configuration variable soit en fournissant de l'information ciblée à l'occupant ou au gestionnaire. Le système de pilotage optimal repose sur des mesures recueillies in situ, sur des algorithmes de calibrage temps réel de modèles (observateur d'état), des algorithmes de contrôle auto-adaptatif, des actionneurs pour les systèmes pilotés à distance et une interface homme-machine pour le retour d'information à l'utilisateur.

Mes travaux de thèses s'inscrivent dans ce projet et se focalisent sur l'identification des modèles thermiques de bâtiments. En effet, si l'on souhaite minimiser la consommation énergétique d'un habitat, système inertiel, la gestion anticipative reste la meilleure solution. Elle permet d'anticiper les changements futurs au niveau des contraintes et des sollicitations.

Cette gestion anticipative nécessite la modélisation thermique dynamique du bâtiment, afin d'en prédire la température intérieure.

I.2 La gestion anticipative

Un modèle du bâtiment est nécessaire dès lors que l'on veut anticiper la température intérieure d'un bâtiment en vu de sa gestion. Actuellement, il existe différentes approches de gestion. Les approches sans modèles présentant des limites, comme il sera vu dans la première partie, nous nous intéresserons à la modélisation dynamique des bâtiments, qui permette d'obtenir de meilleurs résultats de gestion.

I.2.1 Les approches sans modèle et leurs limites

La régulation thermique peut-être réalisée sans modèle de bâtiment, à l'aide d'un thermostat physique. Il s'agit d'un simple régulateur à hystérésis qui déclenche un relais si la température intérieure dépasse un seuil (comparateur AOP). Les différents acteurs qui interviennent par rapport à la régulation thermique sont les suivants :

- les **occupants** : ils choisissent la température de consigne qui correspond à leur confort thermique ;
- les **systèmes de chauffage** : qui chauffent la pièce ;
- les **systèmes de gestion** : qui pilotent le fonctionnement du chauffage de manière à coller à la température de consigne.

Les apports de chaleur par chauffage sont ajustés pour compléter les apports internes et solaires en fonction des conditions météorologiques. Mais cette régulation ne fonctionne que si les changements d'ordre thermique ne sont pas trop importants. De plus, le seul objectif de cette régulation est le maintien du confort thermique. Il ne peut y avoir d'objectifs complémentaires comme la minimisation de la consommation énergétique ou la qualité de l'air.

Il s'agit donc simplement de garantir une température de consigne. Cette température peut être constante, souvent fixée à 19°C. Ces types de régulation se retrouvent principalement dans les grands bâtiments tertiaires et commerciaux. Ils sont généralement appelés BMS (Building Management Systems).

On trouve également des programmeurs avec des températures de consigne variable dans les bâtiments occupés par intermittence. Ceci requiert un contrôleur pour piloter le système de chauffage. Il s'agit de suivre les variations des températures de consigne en fonction des moments d'occupation. Des approximations peuvent être réalisées pour simplifier les calculs. Ces simplifications peuvent induire des erreurs de calcul, comme on peut le voir dans les expériences faites en chapitre 3 de la thèse de Ion Hazyuk [19].

En pratique, il existe deux solutions qui sont exploitées lorsque la température de consigne change de valeur (lorsqu'on passe d'un moment où le bâtiment est inoccupé à un moment d'occupation) :

- Préchauffage avant occupation : le basculement de consigne se fait avant l'arrivée des occupants. Cela nécessite de connaître l'inertie du système.
- Changement de consigne graduel : la température suit une rampe pour passer de 17°C à 19°C entre 2h et 8h par exemple.

Bien que limitant la consommation énergétique en comparaison d'un habitat non équipé de thermostat, ce genre de régulation trouve ses limites dès qu'il s'agit de faire de réelles économies d'énergie.

I.2.2 Les approches à base de modèles

Lorsque l'on veut faire des économies d'énergie, on cherche à atteindre les meilleurs compromis coût/confort en tenant compte du contexte, notamment des apports énergétiques. Or la plupart du temps, ces apports ne sont pas disponibles au moment où ils sont le plus utiles. Par exemple, en été, c'est la nuit qu'il fait le plus frais alors que c'est pendant la journée que l'on souhaiterait pouvoir rafraîchir un bâtiment.

Afin de tirer le meilleur parti de ces apports mais également d'éviter le gaspillage énergétique, il est impératif de pouvoir anticiper les conditions météorologiques à horizon 24h, mais également à plus grand horizon (quelques jours, une semaine). Connaître la météo à quelques jours peut modifier les scénarios optimaux. En effet, connaissant un éventuel rafraîchissement dans les prochains jours, on acceptera peut-être de dépasser légèrement la consigne de température pour quelques heures en cherchant à accumuler des apports solaires qui sont gratuits. De même, on peut anticiper l'occupation et commencer à chauffer en hiver pour arriver progressivement à la température de confort lorsque l'occupant arrive. Cette approche de gestion du confort doit être gérée dans une approche multi-critère afin de prendre en compte par exemple des aspects économiques ou environnementaux.

La modélisation dynamique d'un bâtiment permet d'anticiper sa température interne en fonction, entre autres, de la température extérieure et permet donc de gérer au mieux les apports internes et externes en fonction des dispositifs techniques que possède le bâtiment (capacité de stocker la chaleur ou la fraîcheur, sur-ventilation nocturne afin de ne pas dépasser un maximum en température pendant une journée d'été, etc.). Ainsi, on minimise les besoins en chauffage ou en climatisation et donc en énergie.

I.3 Obtention des modèles pour la gestion

I.3.1 Identification des systèmes

I.3.1-i Définitions

Un **système** est un objet à l'intérieur duquel des variables interagissent entre elles et produisent des signaux observables. Ces grandeurs constituent les sorties du système.

Un système est excité par des stimuli externes appelés entrées. Si ces entrées sont manipulables par l'observateur, alors elles sont dit commandables et appelées **commandes** du système. Si elles ne le sont pas, ce sont des **perturbations**.

Un **modèle** permet de relier les entrées et les sorties d'un système à l'aide d'équations. En fonction de la nature de ces équations, le modèle peut être statique (équations algébriques ou aux dérivées partielles) ou bien dynamique (équations différentielles) : l'état de sortie à un moment t dépend de l'entrée au moment t mais également de ses valeurs précédentes.

Il est également caractérisé par son domaine de validité, c'est-à-dire par le domaine de l'espace des entrées dans lequel l'accord entre les valeurs des sorties du processus calculées par le modèle, et leur valeurs mesurées, est considéré comme satisfaisant compte tenu de l'utilisation que l'on fait du modèle.

Un **paramètre** est une valeur numérique qui n'est pas calculé par le modèle et qui n'est pas une variable d'entrée mesurée ou observée. Les paramètres permettent de caractériser un système et de le différencier facilement par rapport à un autre en prenant des valeurs différentes. L'estimation des paramètres consiste à approcher les valeurs de ceux-ci à partir de données expérimentales et/ou d'informations issues de l'expertise.

I.3.1-ii Structure des modèles

En fonction du mode de conception, on trouve plusieurs sortes de modèles.

Les modèles de **connaissance** (aussi appelés modèles "boite-blanche") sont construits à partir d'une analyse physique en appliquant des lois qui régissent les phénomènes intervenant au sein des processus étudiés. Les modèles d'**expérience** ou **universels** (dit modèles "boîte-noire") sont des structures de modèles-types dont les paramètres sont obtenus à partir des mesures d'entrées et de sorties du phénomène étudié. Tandis que les premiers sont obtenus par la connaissance, les modèles boîte-noire ne décrivent que les relations d'entrées-sorties déduites d'expérimentations. Les paramètres utilisés dans ces relations sont identifiés à partir de mesures d'entrées et de sorties disponibles. On qualifie cette identification d'apprentissage. Cependant, les paramètres et les variables internes de ce dernier type de modèle n'ont pas de sens physique et bien qu'ils soient appris à partir des mesures, la qualité du modèle dépend de la richesse spectrale des jeux de données. Ils ont généralement une faible capacité d'extrapolation, c'est-à-dire une faible capacité à représenter en simulation les phénomènes non présents dans les jeux de données servant à l'identification.

A l'inverse, un modèle de connaissance peut extrapoler des comportements mais il arrive fréquemment que le processus à étudier soit trop complexe ou que les phénomènes qui le régissent soient trop mal connus.

En pratique, lorsque des connaissances, exprimables sous formes d'équations, sont disponibles, mais insuffisantes pour concevoir un modèle de connaissance satisfaisant, on peut avoir recours à une modélisation **semi-physique** (souvent appelés "boîte-grise"). Ce type de modélisation prend en considération à la fois les connaissances et les mesures. Une telle

démarche peut concilier les avantages de l'extrapolabilité d'un modèle de connaissance avec la souplesse d'un modèle comportant des paramètres ajustables.

I.3.1-iii Estimation paramétrique et problème inverse

Soit θ un vecteur non accessible à la mesure et y une variable observable, qui lui est liée par des lois physiques $y = f(x, \theta)$. Un **problème inverse** a pour but d'inverser ces lois physiques pour accéder indirectement à θ , connaissant y [21]. Selon Hadamard [18], pour que ce problème mathématique $y = f(\theta)$ soit **bien-posé**, il faut qu'il réponde à trois conditions :

- **Existence** : $\forall x, y \in X \times Y, \exists \theta \in P$;
- **Unicité** : le vecteur des solutions θ est unique dans P ;
- **Continuité** : la dépendance de θ par rapport à x et y est continue, c'est-à-dire que si $dx, dy \rightarrow 0$ alors $d\theta \rightarrow 0$.

Nous noterons toutefois que le fait que le problème mathématique soit bien posé n'implique pas qu'il soit bien conditionné. La continuité est une condition nécessaire, mais pas suffisante, de stabilité de la solution. Ainsi, si le problème est **mal conditionné** [8], alors sa solution sera toujours sensible aux incertitudes des observations. On définit le conditionnement d'un système par une quantité numérique qui mesure la dépendance de la solution d'un problème par rapport aux données du problème.

Soit S un système dynamique à états continus. Un tel système peut-être décrit par les équations d'état suivantes :

$$\begin{aligned}\dot{X} &= f(X, \theta, u) \\ Y &= h(X, \theta) \\ X(0) &= X_0(\theta)\end{aligned}\tag{I.1}$$

avec X le vecteur des variables d'état, u le vecteur d'entrée, Y le vecteur des sorties observées et θ le vecteur des paramètres supposé inconnu. X , Y et u sont toutes les trois des variables fonction du temps.

Il s'agit de choisir une structure de modèle, d'estimer le vecteur des paramètres et de caractériser l'incertitude sur l'estimé à partir des mesures sur les entrées, les sorties et les conditions initiales.

On peut obtenir le vecteur de mesure des variables de sortie du système au moment t_k par l'équation suivante :

$$Y_m(t_k) = Y(t_k, \theta^*) + v(t_k)\tag{I.2}$$

où t_k pour $k = 1, \dots, N$ correspond à l'ensemble des instants de mesure, θ^* au vecteur des paramètres "vrais" du système et $v(t_k)$ à la variable aléatoire regroupant les effets conjugués des bruits de mesure, des perturbations environnantes et des erreurs numériques de simulation.

L'estimation des paramètres du modèle est un problème inverse. Certains facteurs peuvent influencer la qualité de cette inversion :

- La complexité de la structure du modèle ;
- Les conditions expérimentales (les signaux d'entrée u , les conditions initiales X_0 et les instants de mesure) ;
- L'estimateur (fonction qui fait correspondre à une suite d'observations la valeur estimée des paramètres $\hat{\theta}$) ;
- L'adéquation entre la structure du modèle et les données disponibles v .

En terme de gestion thermique du bâtiment, il s'agit d'anticiper la température intérieure. Les flux de chaleur sont les entrées du modèle de bâtiment tandis que la température intérieure est sa sortie. Ainsi, la charge de chauffage est calculée en inversant le modèle et en imposant une température à l'air intérieur.

Il existe de nombreux modèles de diverses formes avec divers degrés de formalisme mathématique. C'est l'utilisation attendue du modèle qui détermine le sens et la complexité du modèle. Ainsi, pour modéliser le bâtiment, il existe des modèles de conception, qui sont extrêmement détaillés, et les modèles réduits, qui sont plus légers. La partie suivante détaille ces deux types de modélisation et leurs intérêts pour la gestion anticipative.

I.3.2 Modélisation thermique des bâtiments

I.3.2-i Les modèles de conception

Un modèle de conception est un modèle thermique détaillé qui permet de caractériser un bâtiment dans son ensemble de manière très précise. Ces modèles s'inscrivent dans une démarche de compréhension physique du comportement du bâtiment. Ils reposent sur une connaissance a priori du système à construire ou à rénover et des scénarios d'excitation.

Il existe de nombreux logiciels de modélisation thermique dynamique du bâtiment. Les principaux utilisés sont les suivants [27] :

- **Pléiades+Comfie** : Simulation dynamique de l'enveloppe du bâtiment mais non prise en compte de la présence d'humidité et une modélisation très partielle de la ventilation.
- **TRNSYS** : Référence en matière de simulation dynamique de la thermique et de l'aéraulique des bâtiments. Ce logiciel est surtout utilisé pour la simulation des systèmes énergétiques dans le bâtiment.
- **EnergyPlus** : Ce logiciel possède de vastes fonctionnalités et offre l'avantage d'être gratuit et en open source, et ainsi avec une grande communauté d'utilisateurs.
- **ESP-r** : Permet de modéliser les transferts de chaleurs, d'air, d'humidité et les flux d'énergie électrique. Il est possible d'afficher des résultats avec des bandes d'incertitudes.
- **Matlab/Simulink/SIMBAD** : Environnement de simulation graphique utilisant Matlab comme solveur. Permet d'étudier la régulation des systèmes (pas de temps très petit possible).

- **Dymola (bibliothèque Building Energy and Control Systems)** : Solveur d'équation utilisant le langage de modélisation standard Modelica.

Ce genre de modélisation est très utile en amont de la construction d'un bâtiment afin d'en simuler le comportement et de dimensionner au mieux l'enveloppe et ses dispositifs de chauffage et de climatisation. En effet, il permet de déterminer les déperditions lorsque le bâtiment est maintenu à une certaine température.

Les modèles de conception permettent de comprendre l'influence de tel paramètre ou de tel dispositif mais ils représentent mal un bâtiment réel. Ils sont conçus a priori et jamais recalés sur des données expérimentales car comportant trop de paramètres.

En général, pour la gestion on préférera les modèles légers et simplifiés qui permettent le recalage.

I.3.2-ii Les modèles réduits

Lorsque l'on parle de réductibilité, on cherche à savoir pourquoi et de combien on peut réduire la dimension d'un modèle sans introduire de perte significative de précision.

Parmi les modèles réduits, on retrouve :

- les modèles d'ordre réduit (réduction numérique d'un modèle physique complexe)
- les modèles thermiques simplifiés (simplification par expertise en posant des hypothèses telles que l'homogénéisation de la température dans les matériaux, la limitation aux principaux flux thermique) : on utilise le formalisme des circuits à constantes localisées (R,C).

I.3.2.ii - a) Les modèles d'ordre réduit

Il s'agit de s'affranchir de la complexité calculatoire des modèles décrits plus haut afin de réduire le nombre de paramètres à estimer. Ils permettent de restreindre les calculs répétitifs, ce qui rend leur utilisation plus aisée, sans pour autant perdre de l'information sur la précision des résultats. Cependant, cette réduction fait perdre la dépendance paramétrique des modèles (ce qui est fondamental en conception, mais moins en gestion).

Pour réduire ces modèles, deux approches peuvent être utilisées (on pourra se référer aux thèses de Petit [32] ou Menezo [26] pour plus de précisions) :

- **La méthode de réduction modale** : Il s'agit d'un changement de base afin de diagonaliser la matrice d'évolution A d'un système d'état ([7] et [16]). Un nombre réduit d'éléments propres suffit à reconstituer la solution recherchée avec une bonne précision. Les valeurs propres, et donc les constantes de temps du système, les plus significatives sont conservées. Toutefois, cette méthode ne conserve pas les modes qui sont les plus sensibles aux sollicitations et qui agissent le plus sur la sortie (notions de commandabilité et d'observabilité). Cette méthode dite de Marshall [24] ne s'appuie que sur un critère temporel (constante de temps) pour effectuer la troncature. Elle peut être améliorée par l'utilisation de la méthode de Moore [28].

- **La méthode de réduction de Moore ou réalisation équilibrée** : Cette méthode consiste à optimiser le placement des pôles et des zéros afin de conserver les parties de l'état les plus commandables et observables. Dans cette méthodes, les pôles et les zéros sont approchés par des systèmes de premier ordre [9] afin d'éviter les oscillations dues à la présence de complexes conjugués. Cette méthode présente des propriétés de convergence aussi intéressantes que la méthode précédente. Cependant, la matrice de transformation n'étant pas orthogonale, la réalisation n'offre pas de découplage dynamique des équations d'état. De plus, elle est calculée en fonction des entrées et des sorties du modèle. Un changement du contenu des vecteurs des sollicitations et/ou des sorties entraîne donc une réalisation équilibrée différente. Enfin, le calcul de cette représentation est bien plus coûteux que celui de la méthode précédente mais plus performant.

I.3.2.ii - b) Les modèles thermiques simplifiés

Une analogie physique de la thermique avec les circuits électriques a été largement utilisée dans la littérature scientifique ([1], [20], [25], [15] et [17]). La thèse de Chadia Zayane présente en particulier l'identification d'un modèle thermique de bâtiment représenté par un circuit électrique à partir d'une courbe de charge [38].

Il s'agit de modéliser le bâtiment par un circuit électrique équivalent simple. L'enveloppe du bâtiment est modélisée par des équations différentielles où le nombre de nœuds thermiques (discrétisation spatiale des équations différentielles partielles) est réduit.

TABLEAU I.1 – Tableau des analogies courantes entre thermique et électricité

| Équivalent électrique | Équivalent thermique |
|-------------------------|-----------------------|
| Courant électrique | Flux de chaleur |
| Potentiel électrique | Température |
| Conductance électrique | Conductance thermique |
| Condensateur électrique | Capacité thermique |

Une analogie avec l'électricité peut être réalisée (cf. tableau I.1). Un tel circuit équivalent est une représentation graphique qui permet de comprendre la structure du modèle par une observation rapide du circuit équivalent. Le circuit est facilement modulable en fonction de la prise en compte ou non de certaines spécificités du bâtiment (pont thermique, prise en compte des portes, des fenêtres, etc.).

Ce genre de représentation est réalisée par un expert. Le nombre d'équations peut être réduit afin de trouver un compromis entre la précision désirée et la taille du modèle. Un ensemble de pièces peut ainsi être représenté par un nœud de température égale à la valeur moyenne des toutes les températures des pièces. Les résistances thermiques de couches en série peuvent être sommées tandis que dans le cas de couches parallèles, on somme l'inverse des résistances. On peut réaliser le même genre de raisonnement concernant les capacités thermiques.

La plupart des outils de simulation thermique dynamique utilisent le même principe pour modéliser les phénomènes de conduction thermique. De plus, le niveau de discrétisation est plus grand (conservation des flux, matériaux discrétisés, capacités séparées) et il existe certaines hypothèses (comme l'unidirectionnalité du flux dans le mur par exemple).

Un bâtiment est un système physique dynamique, il est en interaction continue avec son environnement, ce qui rend très complexe sa modélisation avec une haute finesse. Des approximations, des hypothèses et des simplifications permettent de faciliter cette tâche mais peuvent entraîner des erreurs sur les résultats.

I.4 Problématiques de calibration des modèles

Les retours d'expérience montrent des écarts importants entre les prévisions et les vraies consommations des bâtiments, écarts pouvant aller jusqu'à 200 % d'erreur en prédiction basée uniquement sur la connaissance [34].

I.4.1 Les sources d'incertitudes agissant sur un bâtiment

On trouve des sources d'incertitudes dans diverses phases de vie du bâtiment. On distingue trois phases (Projet ANR Habisol Fiabilité 2014, [27]) :

- La phase de conception
- La phase de construction et de mise en œuvre
- La phase d'exploitation

Pendant la **phase de conception**, des simulations sont réalisées pour prévoir les indicateurs de performance du bâtiment à l'étude. Durant cette phase, les incertitudes peuvent venir des hypothèses de modélisation, des paramètres statiques (paramètres thermophysiques des matériaux de l'enveloppe, caractéristiques de vitrages, débit d'infiltration...) ou peuvent être liées aux sollicitations variables ou paramètres dynamiques (données météo, scénarios d'occupation).

Concernant la **phase de conception et de mise en œuvre**, les différents acteurs de la construction confirment que la qualité d'exécution reste un point très important à améliorer si on veut garantir une performance thermique maîtrisée du bâtiment.

C'est pendant la **phase d'exploitation** qu'apparaissent les incertitudes liées à l'occupation et la modélisation du confort thermique qui ont été fixées lors de la conception. Des tests doivent être réalisés afin de contrôler les performances annoncées lors de la phase de réception. Alors que certains contrôles peuvent être effectués dès réception du bâtiment (présentation, niveaux d'éclairage, mesures acoustiques, etc.), d'autres nécessitent plusieurs années de fonctionnement (saison de chauffe, confort d'été, consommation, etc.). De plus, à plus long terme, les propriétés des matériaux peuvent également évoluer, du fait du vieillissement ou de la dégradation du système constructif.

Ainsi, de nombreuses sources d'incertitudes peuvent affecter le comportement thermique d'un bâtiment durant ses différentes phases. C'est là que réside tout l'intérêt d'un modèle physique simplifié. En effet, ce type de modèle sera plus robuste aux sollicitations et perturbations du bâtiment. De plus, il doit être aisément recalibrable afin d'être réajusté au cours du temps.

I.4.2 Problématique de la thèse : Robustesse en identification et recalibrage des modèles

Il est difficile de faire correspondre la complexité d'un modèle à la richesse des jeux de données disponibles à la mesure. Comment savoir si la valeur des paramètres physiques identifiés correspond à des caractéristiques intrinsèques du bâtiment ou bien seulement à l'émanation d'un phénomène particulier et ponctuel contenu dans le jeu de données utilisé pendant l'identification? Une identification sur un jeu de données d'été donnera-t-elle les mêmes valeurs de paramètres qu'une identification réalisée sur jeu d'hiver?

I.4.2-i Objectifs : obtention de modèles robustes

L'objectif est d'obtenir des modèles coïncidant avec des données expérimentales pour la gestion énergétique (notamment en anticipation) c'est-à-dire à horizon 24h pour une résolution d'1h. La robustesse des modèles face aux incertitudes liées aux mesures ou aux phénomènes (occupation, météo) a donc été étudiée. En effet, la capacité d'extrapolation d'un modèle est très pertinente dans le cadre de la gestion anticipative. De plus, le bâtiment étant en constante évolution, il paraît naturel de recalibrer fréquemment un modèle dédié à la gestion. Il est nécessaire de trouver le juste compromis entre précision en identification et précision en validation.

I.4.2-ii Moyens : analyse des problématiques liées à l'identification

Il existe diverses approches en terme de modélisation des bâtiments. Les modèles universels présentent un certain nombre de limites qui seront listées en chapitre II. Les modèles inspirées de la physique présentent quant à eux certains avantages, comme l'intégration de la connaissance physique, la capacité d'extrapolabilité ou la capacité à être validé par la physique. Une première approche en identification à l'aide d'un modèle physique sera réalisée en chapitre III. Cependant, certains points ne doivent pas être négligés si l'on veut avoir de bons résultats. Ainsi, la complexité du modèle doit être déterminée en fonction de la richesse spectrale des jeux de donnée. De plus, certains phénomènes non maîtrisés peuvent survenir et venir perturbés le phénomène d'estimation paramétrique. Ces problématiques seront détaillées en chapitre IV.

L'identifiabilité des modèles a donc été étudiée afin d'améliorer l'identification des modèles en vue de la gestion anticipative.

I.4.2-iii Solutions : analyse de l'identifiabilité des modèles et pilotage optimal de l'estimation paramétrique

Lorsque l'on réalise une série d'identifications, on peut observer le comportement paramétrique et celui de l'erreur. La dispersion de ces facteurs traduit un problème d'identifiabilité du modèle. De plus, la valeur des paramètres dans l'espace de validité qui définit les zones de recherches paramétrique est également important.

Ainsi, une analyse de la dispersion paramétrique a mené à une réduction du modèle dans un premier temps. Les résultats de cette méthode seront présentés en chapitre [V](#). Ensuite, afin d'éviter aux valeurs des paramètres de sortir des zones de validité physique, l'identification a été pilotée par un système de méta-optimisation. Les résultats de cette méta-optimisation sont exposés en chapitre [VI](#).

CHAPITRE II

Mise en évidence sur un cas d'étude des limites des approches à base de modèle d'observation

Table des matières

| | |
|---|----|
| II.1 Description du cas d'étude : les salles PREDIS MHI | 33 |
| II.1.1 Description générale | 33 |
| II.1.2 Mesures et simulation | 34 |
| II.2 Application d'approches d'identification type séries temporelles | 42 |
| II.2.1 Les modèles universels | 42 |
| II.2.2 Résultats d'identification | 43 |
| II.3 Instabilité des modèles | 51 |
| II.4 Besoin de modèles non-linéaires | 52 |

Les modèles d'observation, aussi appelés modèles universels ou encore boîtes-noires, sont des modèles construits à partir de structures type et d'observations issues du processus à étudier. Les signaux expérimentaux utilisés en entrée de ces modèles permettent d'identifier les paramètres qui minimisent l'écart entre une grandeur simulée et une mesure correspondante.

Étudiés depuis de nombreuses années, on trouve une grande quantité de modèles d'observation dans la littérature (ARX, ARMAX, OE, ...), et ils sont souvent associés à des techniques d'optimisation. Des formes particulières permettent de calculer de manière spécifique et adaptée la solution du problème d'identification ([23]). Les structures des modèles sont choisies pour assurer une optimisation facile. Ceci est un avantage considérable lorsque l'on cherche à modéliser simplement un processus, ce qui n'est pas le cas lorsque l'on réalise une modélisation à l'aide du modèle boîte blanche qui demande une expertise et une bonne connaissance du processus.

Cependant, ce type de modélisation présente quelques limites. En effet, les structures types de modèles peuvent mal correspondre à une réalité avec ses non-linéarités. Ainsi, les modèles d'observation peuvent avoir une faible capacité à prédire des comportements non représentés dans le jeu de données d'apprentissage. Les modèles obtenus sont liés à une simulation avec un pas de temps donné et de bons résultats en prédiction ne sont pas garantis.

A travers le cas d'étude des salles PREDIS MHI, les avantages et les limites de ces approches seront évaluées. Les salles PREDIS MHI sont composées d'un espace bureau et d'une salle de classe situés dans un bâtiment du campus de Grenoble. Il s'agit d'un bâtiment dans un bâtiment équipé de capteurs afin d'être monitoré. L'air est renouvelé par une VMC double flux. L'espace bureau est équipé d'une pompe à chaleur pour compléter les apports de chaleur via un échangeur eau-air connecté à l'air injecté dans les salles. Des mesures de température, de débit d'air et d'énergie électrique sont collectées par des capteurs.

PREDIS MHI a la particularité d'être au standard BBC RT2012 et de ne pas avoir de parois en contact avec l'extérieur, ce qui rend la modélisation particulièrement difficile car il n'y a pas de phénomènes dominants : toutes les parois doivent être considérées.

Afin de proposer les meilleurs scénarios de gestion, la température intérieure doit être anticipée à l'aide d'un modèle simple. Dans ce chapitre, nous avons choisi d'utiliser des modèles d'observation. Ceux-ci ont été testés à l'aide de la toolbox System Identification disponible dans Matlab. Chaque modèle présente des avantages et des inconvénients. Existe-t-il un modèle unique satisfaisant pour toutes les situations ? A travers des exemples simples et concrets, nous avons pu observer les limites des ces modèles non physiques.

Ce chapitre suivra les quatre étapes de l'identification :

1. Acquisition des entrées/sorties sous un protocole d'expérimentation :
Description du cas d'étude et de l'obtention des signaux.
2. Choix ou estimation de la complexité du modèle :
Utilisation de la toolbox Matlab *System Identification*
3. Estimation des paramètres du modèle :
Test de plusieurs structures de modèles universels

4. Validation du modèle identifié :
Analyse des résultats obtenus.

II.1 Description du cas d'étude : les salles PREDIS MHI

II.1.1 Description générale

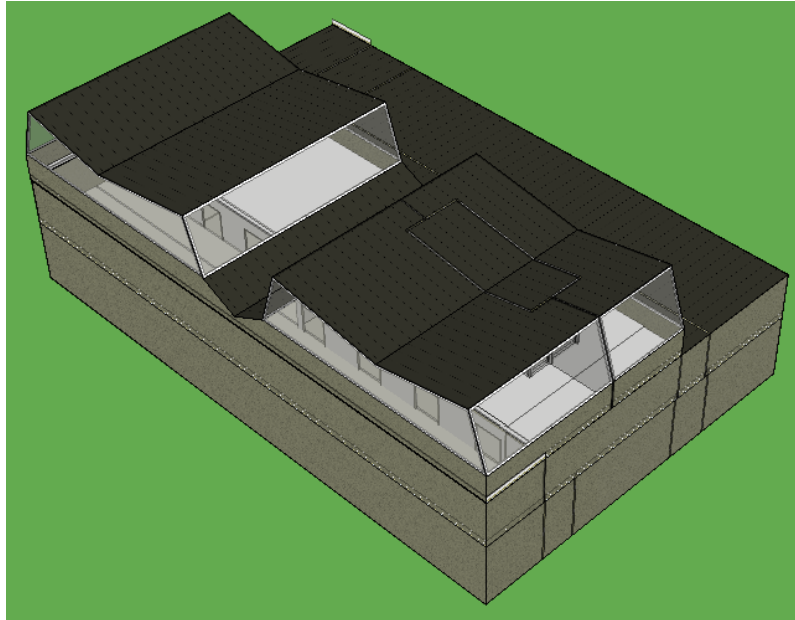


FIGURE II.1 – Vue extérieure du bâtiment PREDIS MHI via le logiciel IDA-ICE

Les salles PREDIS MHI sont situées à l'*Ense*³. Celles que nous allons étudier sont constituées d'un espace bureau et d'une salle de classe. Il s'agit d'un « bâtiment dans un bâtiment ». En effet, ces pièces ont été construites à l'intérieur du bâtiment existant comme on peut le voir sur la vue en coupe de la figure II.2. Aucun mur n'est en contact avec l'extérieur. Les deux pièces possèdent des fenêtres donnant sur un shed, espace vide partiellement vitré qui entoure les salles. La salle de classe possède un shed ou « puit de lumière » : une fenêtre au plafond qui laisse entrer la lumière. Ceci induit un apport de chaleur par flux solaire pour la salle de classe et constitue un apport de lumière important censé limiter la consommation d'énergie électrique pour l'éclairage.

Les murs sont des structures bois isolées avec de la ouate de cellulose tandis que le sol de la salle de classe est constitué d'une épaisse dalle de béton. Quant à l'espace bureau, il y a un espace vide entre le sol et la dalle de béton.

Le renouvellement d'air est réalisé par une VMC double flux. Le chauffage est assuré par une batterie d'eau chaude qui chauffe l'air entrant en hiver. Une pompe à chaleur a été installée dans le bureau au cours de l'été 2014. Tous ces équipements seront décrits plus en détail dans le chapitre III qui traite de la caractérisation des paramètres des appareils électriques.

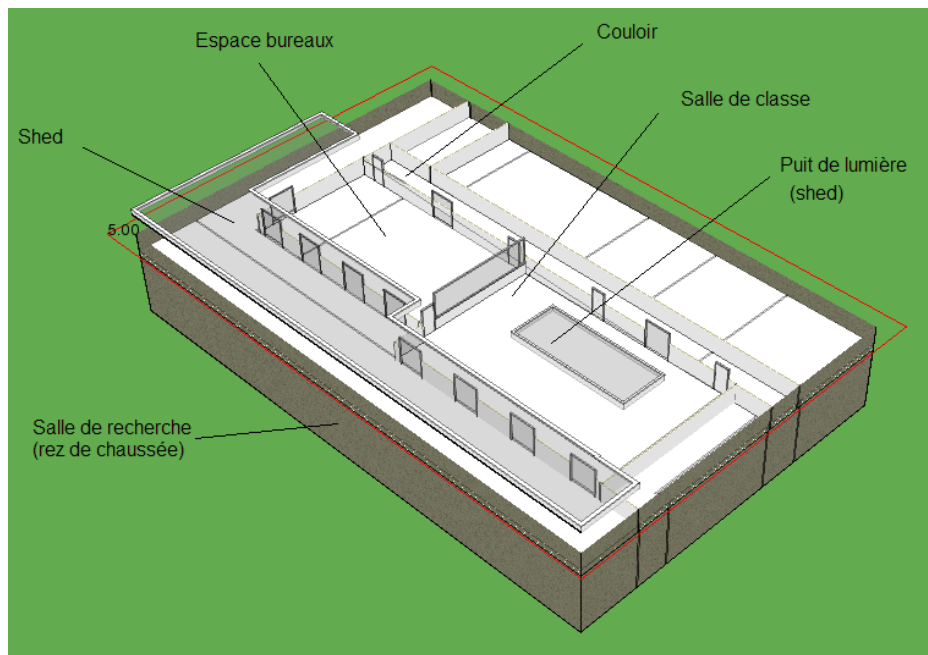


FIGURE II.2 – Vue en coupe de PREDIS MHI via le logiciel IDA-ICE

II.1.2 Mesures et simulation

Les salles PREDIS MHI sont équipées de nombreux capteurs. Les signaux expérimentaux, ou calculés à partir de mesures expérimentales, constituent les entrées des modèles.

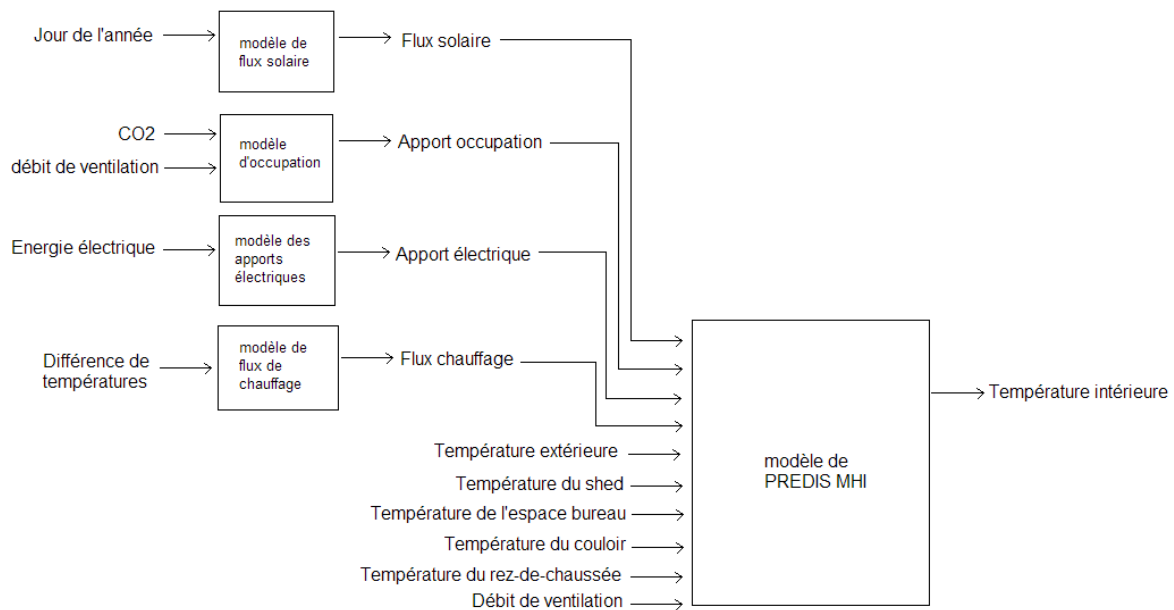


FIGURE II.3 – Structure des modèles de la salle de classe de PREDIS MHI. Bilan des signaux d'entrée nécessaires

Le schéma de la figure II.3 montre toutes les variables que nous utilisons pour la modélisation de la salle de classe PREDIS MHI. Certaines variables sont utilisées directement en entrée du modèle tandis que d'autres sont obtenues par calcul à partir de variables mesurées.

Quand les signaux sont de mauvaise qualité, ont des valeurs manquantes, aberrantes ou quand nous voulons modéliser des situations particulières qui nécessitent de longues périodes de temps, nous avons également recours à des signaux obtenus à partir de logiciel de simulation thermique. Dans ce chapitre, nous avons utilisé des signaux issus du logiciel COMFIE Pléiades où ont été modélisées les salles PREDIS MHI dans le bâtiment complet.

II.1.2-i Signaux d'entrée obtenus par données expérimentales

Les températures sont collectées via des capteurs et utilisées directement en entrée des modèles. Les capteurs utilisés ont une précision à plus ou moins 0,5 °C. Les données enregistrées peuvent contenir des trous ou des valeurs aberrantes. Lorsque ces valeurs erronées sont réparties sur de petits intervalles, on peut facilement les interpoler. Les données contiennent parfois de longues périodes manquantes. La prise en compte et la gestion de ces valeurs manquantes ou aberrantes fait partie de la gestion anticipative.

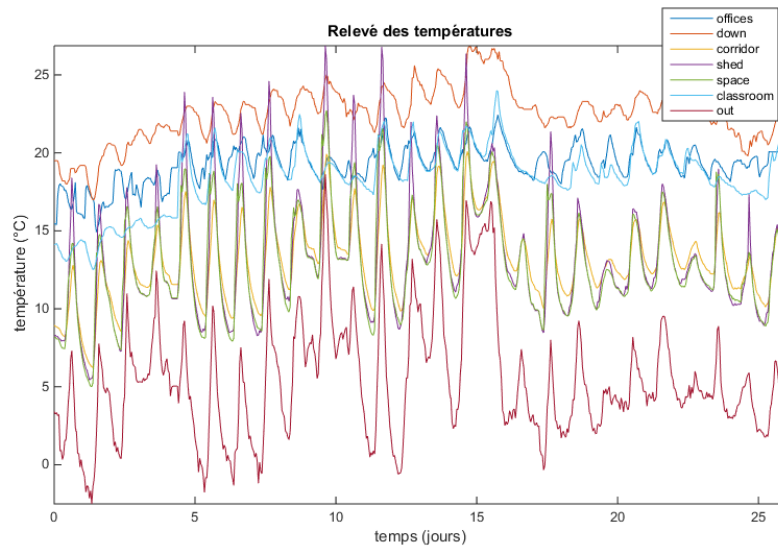


FIGURE II.4 – Relevé des températures des salles PREDIS MHI pour le mois de janvier 2015

La figure II.4 montre le relevé des températures de l'espace bureau et de la salle de classe ainsi que des salles adjacentes et de la température extérieure pour le mois de janvier 2015.

Les apports de chaleurs quant à eux sont obtenus par calcul à partir de mesures expérimentales.

II.1.2-ii Signaux d'entrée obtenus par calcul et données expérimentales

II.1.2.ii - a) Calcul de l'apport thermique dû au chauffage par la VMC double flux

L'air entrant est préchauffé par l'air sortant par contact avec un échangeur rotatif. Le chauffage de l'air se fait par échange avec l'eau chaude de la batterie. La puissance de chauffage a été calculée à l'aide du bilan de chaleur ci-dessous :

$$\rho \times C_p \times Q \times \Delta T = P_{chauff} \quad (\text{II.1})$$

avec ΔT la différence de température amont et aval à la batterie d'eau chaude, ρ la masse volumique de l'eau, C_p la capacité thermique de l'eau, Q le débit d'eau dans la batterie d'eau chaude et P_{chauff} la puissance de chauffage obtenue.

Le débit d'eau a été relevé à intervalles réguliers et pour différentes valeurs de débit d'air. Nous avons remarqué qu'il était globalement constant. Nous avons donc fixé les débits d'eau du bureau et de la salle de classe à des valeurs constantes égales respectivement à $0,0021m^3/h$ et $0,0023m^3/h$.

La figure II.5 montre des profils de chauffage type que nous avons pu obtenir.

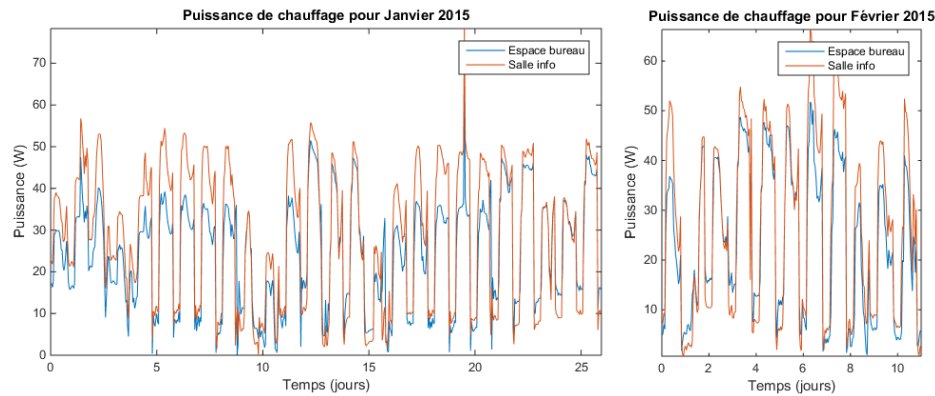


FIGURE II.5 – Puissance de chauffage pour Janvier et Février 2015

II.1.2.ii - b) Calcul de l'apport thermique dû aux occupants et au flux d'énergie électrique

La consommation des appareils, tels que l'informatique ou l'éclairage, est relevée sur les compteurs électriques. La figure II.6 rend compte du type de courbe que l'on peut obtenir. Les signaux sont en escalier avec des fronts montants à chaque nouveau kWh consommé. Ceci rend le calcul de la puissance équivalente difficile car cela génère des pics de puissance lorsque nous dérivons l'énergie.

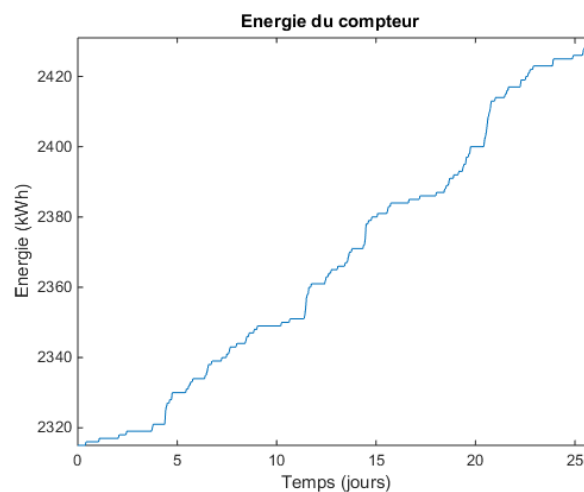


FIGURE II.6 – Energie d'un compteur électrique pour janvier 2015

On peut distinguer l'énergie consommée par les ordinateurs et pour l'éclairage car les compteurs sont dissociés. Si on dérive l'énergie afin de visualiser la puissance, on obtient les courbes de la figure II.7.

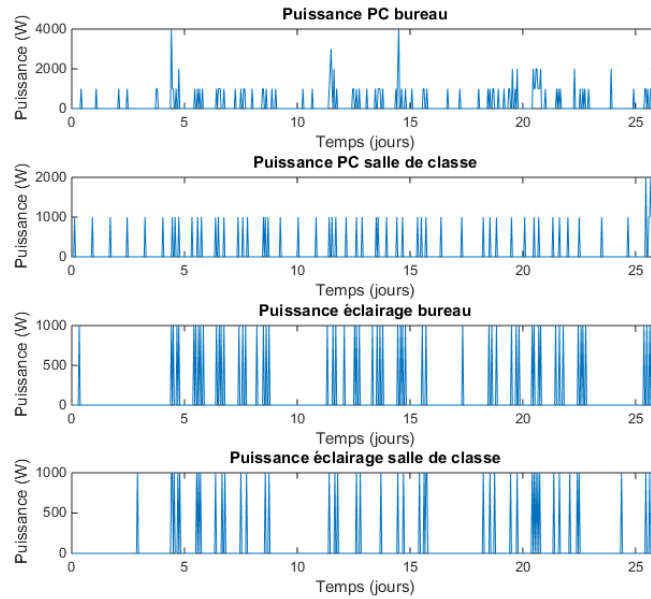


FIGURE II.7 – Puissance obtenue par dérivation de l'énergie électrique

Ce type de signal est inexploitable en entrée d'un modèle car les pics viendraient perturber l'identification. C'est pourquoi nous avons réalisé une moyenne des données interpolées sur un pas horaire. Les courbes obtenues sont observables sur la figure II.8.

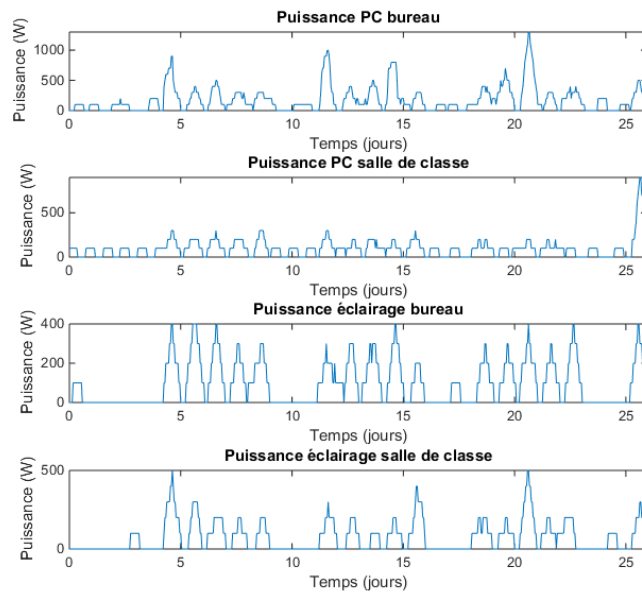


FIGURE II.8 – Puissance obtenue par moyenne glissante à partir de la dérivée de l'énergie électrique

L'apport de chaleur dû à l'occupation peut-être calculé en fonction de la concentration de CO_2 liée au débit de ventilation à l'aide de l'équation ci-dessous :

$$N_p = \frac{\frac{\Delta CO_2}{\Delta t} V + Q_{soufflage}(CO_2 - CO_2^{air})}{Q_{personne} CO_2^{personne}} \quad (II.2)$$

où N_p correspond au nombre de personnes, ΔCO_2 à la différence de concentration de CO_2 entre le point calculé et le point précédent (*ppm*), Δt le pas de temps (*h*), V le volume de la pièce, $Q_{soufflage}$ le débit d'air de la pièce (m^3/h), CO_2 la concentration de CO_2 de l'air sortant, CO_2^{air} la concentration de CO_2 de l'air entrant, $Q_{personne}$ le taux moyen de respiration d'une personne et $CO_2^{personne}$ le taux moyen de CO_2 par personne.

Lorsque la concentration de CO_2 n'est pas disponible (valeurs manquantes ou erronées), on peut avoir recours à la consommation électrique des PC qui rend compte de l'occupation. Cette puissance électrique peut alors être majorée par un coefficient multiplicateur supérieur à 1 afin de représenter l'apport de chaleur dû au flux électrique et l'apport de chaleur dû à l'occupation. La puissance de l'éclairage électrique est à prendre en considération à part puisque le nombre d'occupant n'est pas corrélé à la consommation électrique pour l'éclairage.

$$\begin{aligned} Puissance_{occupation} &= Puissance_{ordinateurs} * coef \\ Puissance_{electrique} &= Puissance_{eclairage} \end{aligned} \quad (II.3)$$

avec *coef* un coefficient multiplicateur supérieur à 1.

II.1.2.ii - c) Calcul de l'apport solaire

Ce paragraphe est inspiré des travaux de Camille Latrémolière dans le cadre d'un stage de deuxième année à propos de la mise au point d'un modèle d'éclairage naturel intégré dans le gestionnaire énergétique du bâtiment Canopea de la compétition Solar Decathlon Europe 2012 [22], [12].

La terre est en orbite autour du soleil selon une trajectoire elliptique. L'inclinaison de l'axe terrestre, l'endroit sur la planète et le jour de l'année sont des critères qui influent sur le temps d'ensoleillement et la puissance transmise.

Pour repérer la position du soleil par rapport aux 4 points cardinaux, il faut calculer l'altitude et l'azimut. L'altitude est l'angle formé par le plan horizontal du lieu considéré et le vecteur qui relie le point local avec le soleil. L'azimut est l'angle horizontal formé par le plan méridien (axe nord-sud) et le plan vertical du vecteur qui relie le point local et le soleil.

Le rayonnement solaire peut être divisé en trois composantes :

- Le **rayonnement direct** qui en constitue la majeure partie ;
- le **rayonnement diffus** qui dépend de l'atmosphère et de sa composition ;
- le **rayonnement réfléchi** par le sol appelé aussi albédo terrestre.

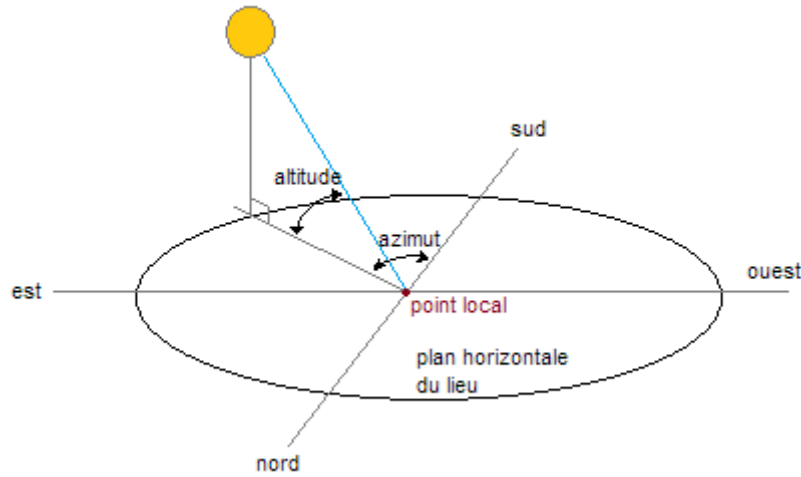


FIGURE II.9 – Caractérisation de la course du soleil par rapport à un point local

Le rayonnement global, somme de ces trois composantes, est ainsi déterminé par trois groupes de facteurs : les relations géométriques entre Soleil et surface terrestre, l'atténuation due à la traversée de l'atmosphère et les facteurs topographiques.

En fonction de la localisation sur Terre et du moment choisi, la position du soleil peut être obtenue par des formules astronomiques. Le facteur d'atténuation atmosphérique quand à lui est conditionné par les particules présentes dans l'atmosphère et dont la quantité varie avec l'épaisseur de cette dernière. Les facteurs topographiques induisent de fortes variations à l'échelle locale. L'inclinaison et l'exposition d'une surface modifient significativement l'angle d'incidence du rayonnement solaire. De plus, des effets de masque dus à des versants opposés ou encore à des avancées sur une fenêtre peuvent également altérer l'intensité du flux solaire.

- Rayonnement direct

Il est exprimé en W/m^2 et en traversant l'atmosphère subit une atténuation de son intensité :

$$R_{direct} = \tau^M \times \Phi_{atm} \times e^{M \times E_{rayleigh} \times L} \quad (II.4)$$

Où R_{direct} représente le rayonnement solaire direct, τ_M le coefficient de transmissivité, Φ_{atm} le flux solaire en sortie d'atmosphère, M la masse d'air optique relative, $E_{rayleigh}$ l'épaisseur optique de Rayleigh et L le facteur de trouble de Linke.

Ces coefficients dépendent du jour de l'année, de la pression atmosphérique, de la température locale et de l'humidité.

L'intensité du rayonnement reçue par une surface inclinée et ayant une certaine exposition est inférieure à celle du rayonnement direct. Il est nécessaire de prendre en compte l'angle

d'incidence entre le rayon solaire et la surface étudiée qui varie en fonction de la position du soleil :

$$Angle = \cos^{-1}(\cos(Altitude)\sin(inclinaison)\cos(Azimet-Exposition)+\sin(Altitude)\cos(Inclinaison)) \quad (II.5)$$

Finalement, la puissance directe reçue est égale au produit de l'angle d'incidence et du rayonnement direct :

$$Puissance = Angle \times R_{direct} \quad (II.6)$$

- *Rayonnement diffus*

Il est exprimé en W/m^2 et arrive sur le plan récepteur après avoir été diffusé par les particules présentes dans l'atmosphère. Il est supposé isotrope (sans direction prédominante), il ne dépend donc que de l'altitude du soleil :

$$R_{diffus} = C_{solaire}(0,271 - 0,294\tau^M)\sin(Altitude) \quad (II.7)$$

- *Rayonnement réfléchi*

Également exprimé en W/m^2 , il arrive sur le plan récepteur après avoir été réfléchi par le sol. Il est lui aussi supposé isotrope, l'orientation du plan récepteur n'a donc pas d'importance, seule son inclinaison en a.

$$R_{réfléchi} = 0,2 + 1367(0,271 + 0,706\tau^M) + \sin(Altitude)\sin(\frac{Inclinaison}{2})^2 \quad (II.8)$$

- *Coefficient d'atténuation due à la couverture nuageuse*

Les données météorologiques classiques fournissent une estimation du rayonnement solaire global mesuré au sol sur une surface horizontale orientée sud. Cette valeur tenant compte de la couverture nuageuse, elle correspond au rapport entre l'estimation météorologique et la valeur théorique du rayonnement solaire sur une surface horizontale orientée sud.

- *Flux solaire des salles PREDIS MHI*

Le gabarit du puits de lumière influe sur le flux solaire entrant dans la salle de classe. Dans l'espace bureau, l'absence de puits de lumières ainsi que l'orientation des fenêtres donnant sur le shed font que le flux solaire est négligeable dans cette zone.

Nous ne disposons pas de capteurs pour quantifier le flux solaire. Celui que nous utilisons en entrée du modèle est obtenu par simulation à l'aide des équations décrites précédemment. Le flux solaire dépend de la date, de la position sur terre, de la nébulosité, de la pression

atmosphérique et d'un niveau de pollution dont les valeurs numériques retenues sont les suivantes :

$$\begin{aligned}
 latitude &= 45,183 \\
 niveau_{mer} &= 330 \\
 nébulosité &= 2 \\
 pollution &= 0,1 \\
 humidité &= 0,2 \\
 albedo &= 0,1 \\
 Facteur_{solaire} &= 0,3
 \end{aligned}
 \tag{II.9}$$

Sur la figure II.10, on peut visualiser le type de courbe que nous pouvons obtenir à l'aide de cet algorithme.

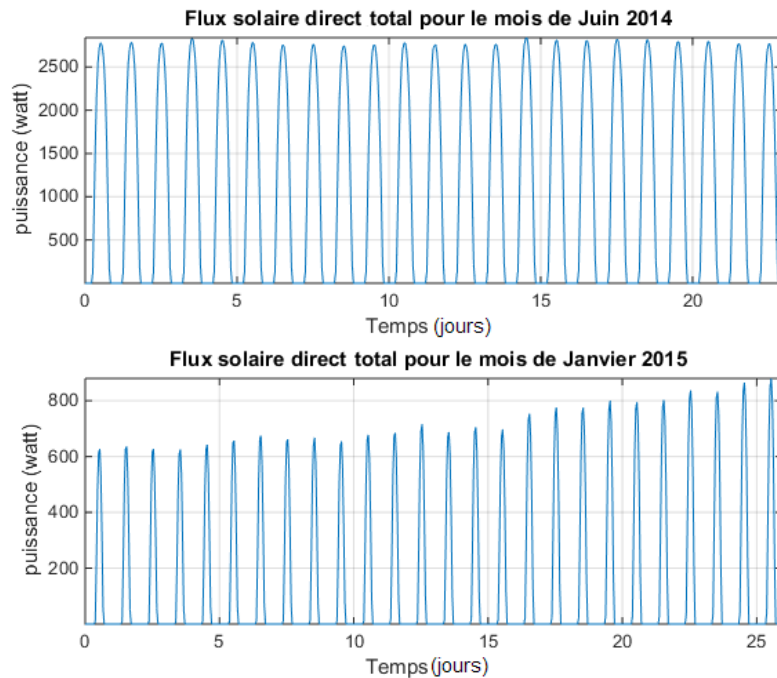


FIGURE II.10 – Flux solaire direct total pour deux jeux de données été / hiver

II.1.2-iii Signaux obtenus par simulation à l'aide du logiciel COMFIE Pléiade

Comfie est un logiciel permettant de réaliser des simulations de l'enveloppe du bâtiment. **Pléiade** est une interface qui permet un apprentissage rapide et limitant le risque d'erreurs lors de la phrase de description du projet.

Les salles PREDIS MHI ont été modélisées dans ce logiciel. Nous avons à disposition un an de données simulées selon des scénarios de ventilation et d'occupation simplifiés.

La salle de classe est occupée entre 8h et 18h par 26 personnes tous les jours de l'année sauf weekend, jours fériés et vacances universitaires. Dès qu'il y a occupation, la ventilation est en marche et le débit est constant et fixé à $420m^3/h$.

L'espace bureau, quant à lui, est occupé de 8h à 18h par 6 personnes tous les jours sauf weekend, jours fériés et jours de fermeture du laboratoire (cela correspond à 4 semaines dans l'année). De même que dans la salle de classe, lorsqu'il y a occupation, le débit de ventilation est constant et fixé à $120m^3/h$. Ces valeurs représentent des scénarios standards simplifiés qui permettent dans un premier temps de mieux appréhender l'estimation paramétrique.

Dans le logiciel COMFIE, la simulation de fenêtres en intérieur n'est pas possible. Ainsi, le flux solaire n'est pas considéré avec ces jeux de données obtenus par simulation. En effet, dans la réalité, le flux solaire traverse le shed, et arrive sur la fenêtre du plafond de la salle de classe.

II.2 Application d'approches d'identification type séries temporelles

Lorsque l'on cherche à modéliser mathématiquement un système, on s'intéresse à la détermination de sa structure et à la détermination des valeurs numériques de ses paramètres.

On peut emprunter deux chemins pour y parvenir :

- On peut partir des lois physiques qui régissent le système. Ceci est pertinent si ces lois sont connues avec assez de précision ;
- On peut partir des données expérimentales et chercher un modèle qui traduit correctement ces données.

La modélisation d'un processus à l'aide des données expérimentales permet d'éviter la modélisation par des modèles physiques purement théoriques qui sont en général difficiles à obtenir.

Pour ce genre de modélisation, on utilise des modèles dit "universels" ou encore "boîtes-noires".

II.2.1 Les modèles universels

Les modèles universels présentent plusieurs avantages. Leur structure est adaptée à l'estimation paramétrique et ayant été étudiés depuis de nombreuses années, il existe nombre d'outils disponibles (régression, réseaux de neurones, Fourier, ondelettes, théorie du chaos). De plus, ces modèles sont souvent très rapides à converger en estimation paramétrique et c'est important pour du temps réel.

Il s'agit, à partir de données expérimentales, d'obtenir le modèle mathématique d'un système. Les séries temporelles (suite de mesures chronologiques d'un processus) montrent des tendances et des schémas répétitifs qui peuvent être appris par des méthodes statistiques.

Parmi les modèles les plus connus, on retrouve :

- Le modèle ARX : il s'agit d'un modèle régressif qui inclut des entrées et un bruit blanc de moyenne nulle.

- Le modèle ARMAX : ce modèle reprend les attributs du modèle ARX mais inclut une fonction de transfert avec une moyenne ajustable sur le bruit blanc. Le bruit blanc permet en général de modéliser des perturbations non-mesurables dans le modèle, or, ces perturbations sont rarement de moyenne nulle et peuvent aussi répondre à un modèle.
- La modélisation par système d'état : il met en relation les variables et leurs dérivées temporelles (appelés variable d'état).

II.2.2 Résultats d'identification

A l'aide de modèles universels, la salle de classe de PREDIS MHI a été identifiée. L'identification de l'espace bureau pourrait être obtenue par analogie. Pour réaliser ces identifications, des jeux de données obtenus par simulation dans le logiciel COMFIE Pléiades ont été utilisés dans un premier temps. Ceci permet de visualiser comment les modèles réagissent aux scénarios standards, entièrement maîtrisés et prédéfinis. Ensuite, des jeux de données expérimentaux ont également été utilisés afin de corroborer les premiers résultats.

La toolbox Ident de Matlab a été utilisée pour identifier les paramètres de plusieurs structures de modèles sur les jeux de données. Il s'agit d'observer dans quelle mesure les modèles universels qui ne possèdent pas de paramètres physiques peuvent représenter la réalité. La plupart de ces modèles sont polynomiaux et les coefficients de ces polynômes sont les paramètres du modèle à identifier. L'espace de recherche n'est pas limité et les coefficients ne possèdent pas d'interactions entre eux. Nous pouvons nous demander dans quelle mesure une identification reflétera un comportement général du processus ou bien simplement une situation particulière liée au jeu de données.

II.2.2-i Résultats d'identification sur les signaux obtenus par simulation

Dans un premier temps, les paramètres du modèle de la salle de classe PREDIS MHI ont été estimés à partir de simulations issues du logiciel COMFIE Pléiades où a été représenté le bâtiment.

II.2.2.i - a) Description des données du modèle

Nous avons à disposition un an de signaux simulés. Les figures II.11 et II.12 correspondent aux courbes des données des mois de Janvier et de Février. Nous remarquons qu'il n'y a plus de ventilation à la fin du jeu de Février. Cela correspond aux deux semaines de vacances universitaires où il n'y a pas d'occupation et donc pas de ventilation.

Comparons maintenant les résultats obtenus durant la phase d'identification et la phase de validation lorsque survient un scénario de ventilation différent et non représenté dans le jeu d'identification.

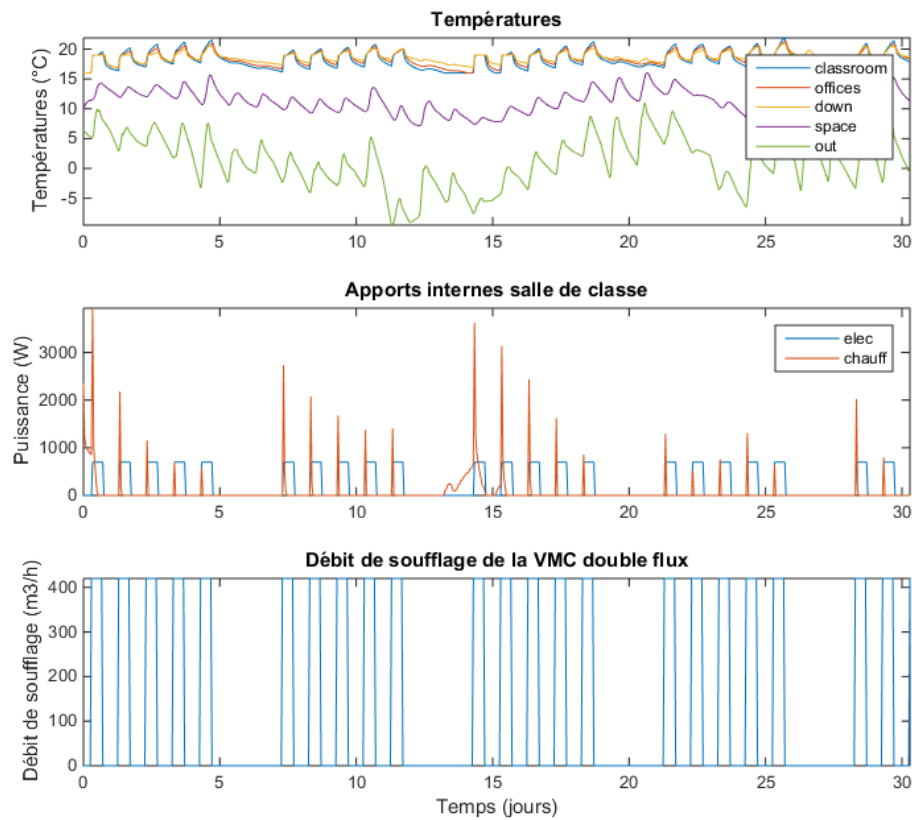


FIGURE II.11 – Signaux obtenus par le logiciel Comfie pour le mois de Janvier

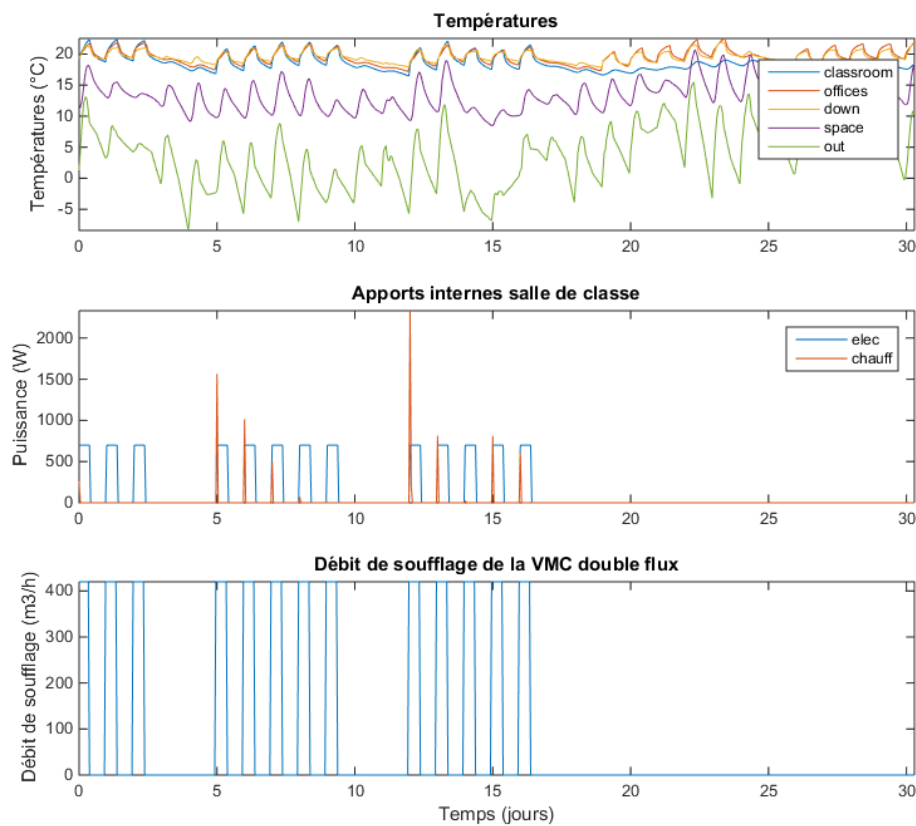


FIGURE II.12 – Signaux obtenus par le logiciel Comfie pour le mois de Février

II.2.2.i - b) Identification de signaux obtenus par simulation avec des modèles universels

- Identification avec un système d'état

Le modèle par système d'état s'écrit sous cette forme :

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) + Ke(t) \\ y(t) &= Cx(t) + Du(t) + e(t)\end{aligned}\tag{II.10}$$

où A , B , C et D sont les matrices du systèmes d'état, K , la matrice des perturbations, $u(t)$ correspond au vecteur des entrées, $y(t)$ à la sortie, $x(t)$ au vecteur des états et $e(t)$ aux perturbations.

Pendant de nombreuses années, le choix de la complexité se faisait par une procédure essais/erreurs. Des progrès significatifs dans ce domaine ont été effectués depuis. En effet, de nombreux algorithmes fiables estiment la complexité des modèles à partir des données [13] et [35]. L'ordre du modèle doit minimiser une mesure de la vraisemblance pénalisée où la pénalité est déterminée par le critère adopté.

— Le critère d'information d'Akaike (AIC) :

$$AIC = 2k - 2\ln(L)\tag{II.11}$$

avec k le nombre de paramètres et L le maximum de la fonction de vraisemblance du modèle. Ce critère à minimiser vise à trouver un compromis entre la qualité de l'ajustement et la complexité du modèle, en pénalisant les modèles ayant un trop grand nombre de paramètres.

— Le critère d'information Bayésien (BIC) :

$$BIC = -2\ln(L) + k\ln(n)\tag{II.12}$$

où n est le nombre d'observations dans l'échantillon étudié. L' AIC pénalise le nombre de paramètres moins fortement que le BIC .

Dans la toolbox *System Identification* de Matlab, c'est la méthode des valeurs singulières de Hankel qui est utilisée dans le cas d'une modélisation par système d'état. Ici, le modèle proposé est d'ordre 3.

Le jeu de Janvier servira à l'identification des paramètres du modèle tandis que le jeu de Février servira à la validation. La tendance moyenne a été retirée des signaux d'entrée et de sortie, c'est pourquoi les courbes de la figure II.13 possèdent une échelle verticale centrée en zéro. Le tableau II.1 présente les valeurs d'erreurs obtenues pendant les phases d'identification et de validation.

Tandis que l'identification (courbes de gauche de la figure II.13) donne de très bons résultats avec une erreur moyenne très faible (0.1741 °C) et une erreur maximale très correcte (1.1202 °C) pour le jeu d'apprentissage, on remarque néanmoins un décrochement pendant

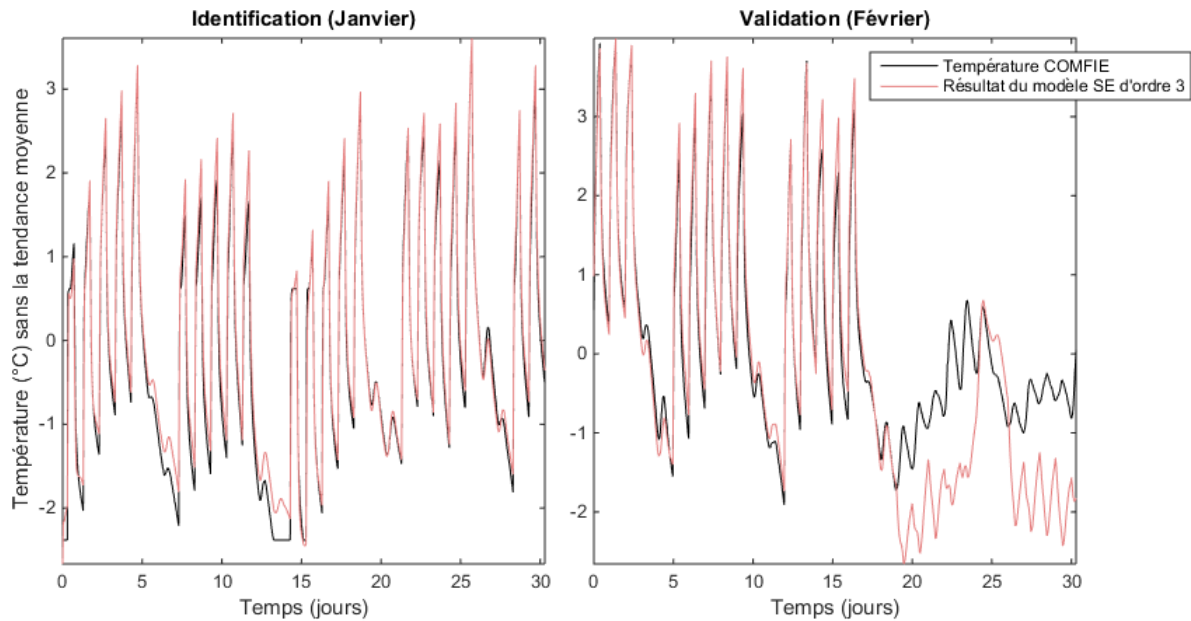


FIGURE II.13 – Identification et validation avec un jeu de données obtenu par simulation avec un système d'état d'ordre 3

| | Identification (Janvier) | Validation (Février) |
|---------------------|--------------------------|----------------------|
| Erreur moyenne (°C) | 0.1741 | 0.5294 |
| Erreur max (°C) | 1.1202 | 2.3082 |

TABLEAU II.1 – Erreurs d'apprentissage et de validation pour le système d'état d'ordre 3 avec des jeux issus de simulations COMFIE-Pléiades

la validation à partir du 17^{ème} jour où survient le changement de scénario. Pendant cette période, il n'y a ni occupation, ni ventilation. Bien que l'erreur moyenne reste bonne (0.5294 °C), ces résultats montrent un problème d'extrapolation. En effet, les situations non comprises dans le jeu d'apprentissage (périodes de vacances) ne sont pas simulées correctement dans le jeu de validation.

Testons maintenant le modèle ARX afin de voir si nous observons de meilleurs résultats.

- Identification avec un système ARX

La forme du modèle ARX est la suivante :

$$y(t) + a_1y(t-1) + \dots + a_nay(t-n_a) = b_1u(t-n_k) + \dots + b_nbu(t-n_b-n_k+1) + e(t) \quad (\text{II.13})$$

où :

$y(t)$ est la sortie à l'instant t ,

n_a le nombre de pôles,

n_b le nombre de zéros plus 1,

n_k le nombre d'échantillons d'entrée qui interviennent avant que les entrées ne perturbent la sortie,

$y(t-1) \dots y(t-n_a)$ les sorties des pas précédents dont dépend la sortie actuelle,
 $u(t-n_k) \dots u(t-n_k-n_b+1)$ les entrées précédentes et retardées dont dépend la sortie actuelle
et $e(t)$ la valeur de la perturbation bruit blanc.

Ici, c'est le critère d'Akaike (AIC) qui est utilisé pour obtenir le meilleur ordre.

Les ordres obtenus avec ce critère sont les suivants :

$$\begin{aligned} n_a &= 1 \\ n_b &= [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1] \\ n_k &= [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1] \end{aligned} \quad (\text{II.14})$$

Puisqu'il n'y a qu'une seule sortie, n_a est de dimension 1. Les n_b et n_k sont de dimensions égales au nombre d'entrées soit 9 entrées pour le modèle COMFIE.

Les résultats d'identification et de validation sont observables sur la figure II.14. Le tableau II.2 donne les valeurs d'erreurs obtenues en identification et en validation.

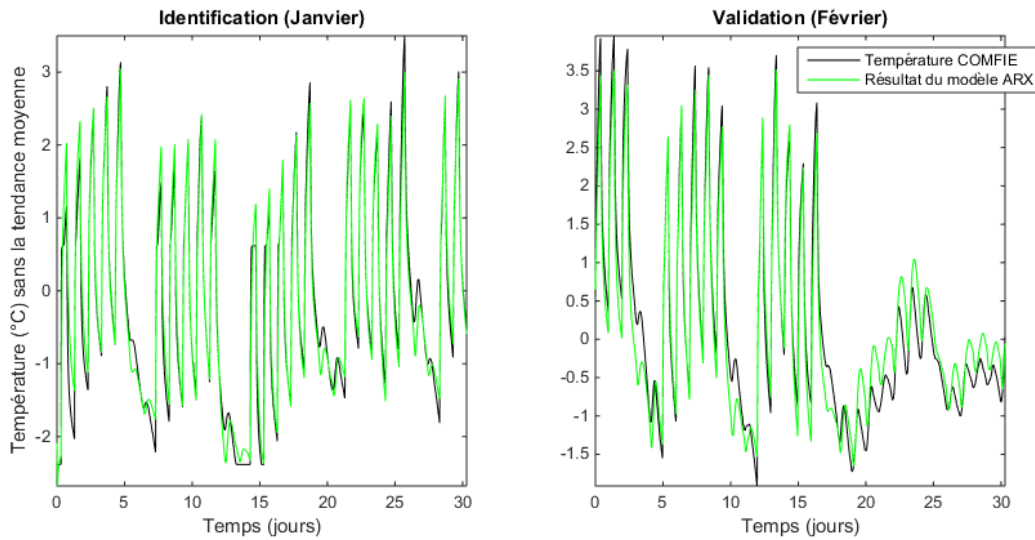


FIGURE II.14 – Identification et validation avec un jeu de données obtenu par simulation avec un modèle ARX

| | Identification (Janvier) | Validation (Février) |
|---------------------|--------------------------|----------------------|
| Erreur moyenne (°C) | 0.2016 | 0.3052 |
| Erreur max (°C) | 1.4662 | 0.8595 |

TABLEAU II.2 – Erreurs d'apprentissage et de validation pour le modèle ARX avec des jeux issus de simulation COMFIE-Pléiades

Le résultat est meilleur que précédemment. En effet, l'erreur maximum en prédiction n'excède pas le degré. Ainsi, nous pouvons remarquer ici que le modèle ARX arrive à extrapoler le comportement de la ventilation malgré l'absence de représentation de cette situation (absence de ventilation prolongé pendant les vacances) dans le jeu de données d'identification. Ce n'est pas le cas du modèle d'état d'ordre 3 qui décroche au moment de la période de vacances.

Nous allons maintenant tester les modèles universels avec des mesures expérimentales.

II.2.2-ii Résultat sur des mesures expérimentales

II.2.2.ii - a) Description du modèle

Nous utiliserons ici des signaux expérimentaux qui ont été collectés par des capteurs.

Aux entrées utilisées précédemment et issues du logiciel COMFIE-Pléiade, il faut ajouter les flux solaires (qui ne sont pas pris en compte dans la modélisation des salles PREDIS MHI dans COMFIE-Pléiade). Nous avons distingué quatre flux solaires : le flux direct est, le flux direct ouest, le flux direct sud et le flux diffus. Ceci permet une plus grande liberté quant à l'impact de chaque flux sur la température intérieure. De même que chaque apport interne a été considéré séparément afin de donner au modèle un plus grand degré de liberté.

Ainsi, le modèle réel que nous avons identifié avec des modèles universels possède 13 entrées.

Avec des données expérimentales, le débit de ventilation et les apports internes sont beaucoup moins réguliers et donc plus difficiles à anticiper. En revanche, les données d'excitation sont plus riches et devraient donc donner de meilleurs résultats.

Les figures II.15 et II.16 montrent deux jeux de données enregistrés en Janvier et Février 2015.

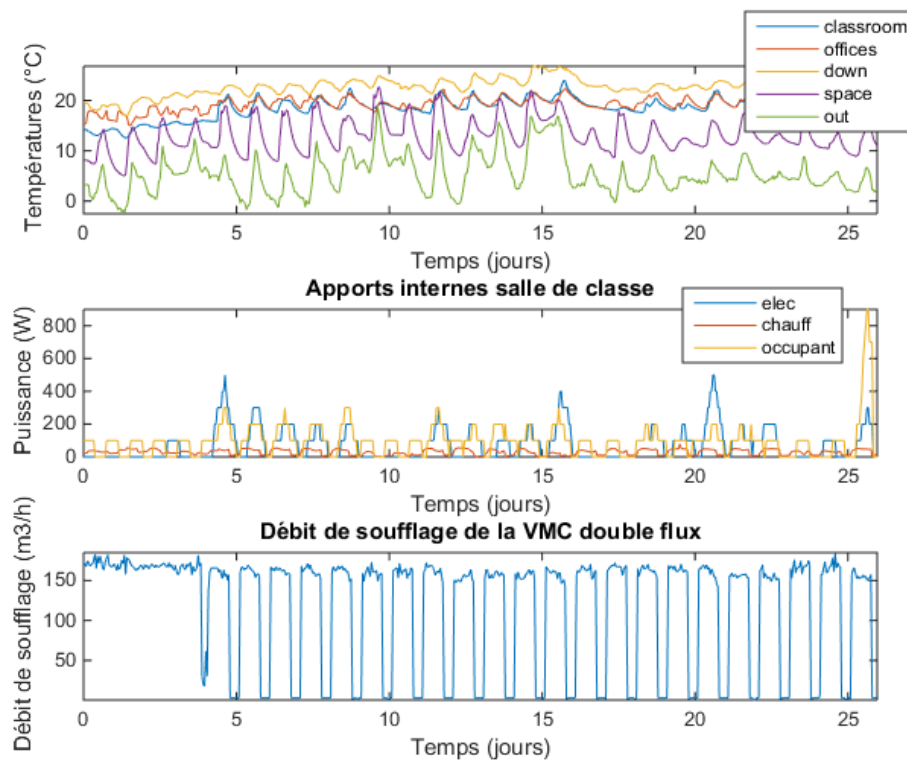


FIGURE II.15 – Signaux d'entrée expérimentaux relevés en Janvier 2015

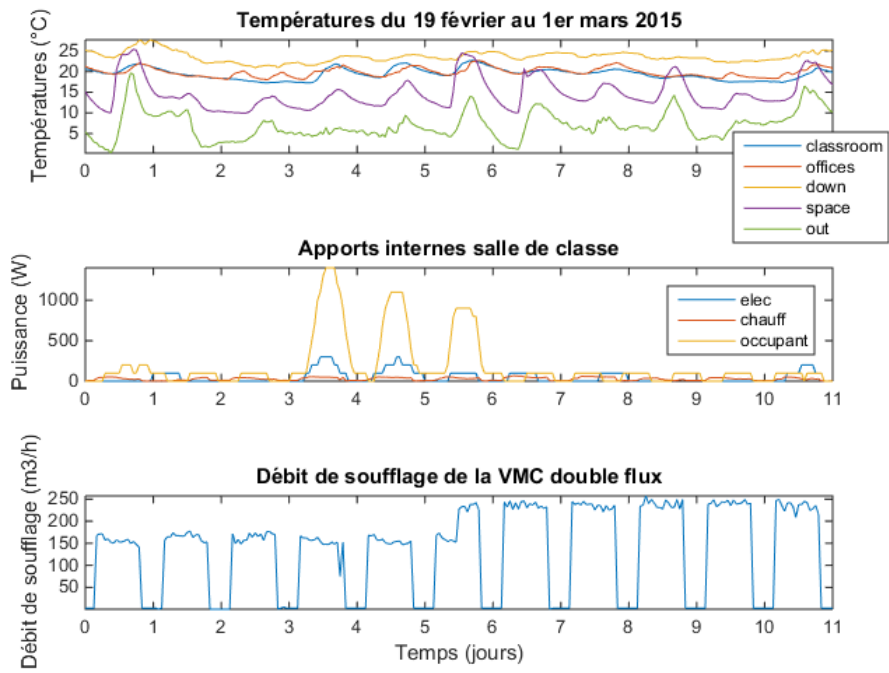


FIGURE II.16 – Signaux d’entrée expérimentaux relevés en Février 2015

II.2.2.ii - b) Identification de signaux expérimentaux avec des modèles universels

- Identification avec un système d’état

Cette fois-ci, l’algorithme propose un ordre 5 pour identifier le modèle en se focalisant sur la prédiction. Les résultats sont observables sur la figure II.17. Les valeurs d’erreurs obtenues sont données dans le tableau II.3.

| | Identification (Janvier) | Validation (Février) |
|---------------------|--------------------------|----------------------|
| Erreur moyenne (°C) | 1.4002 | 4.1063 |
| Erreur max (°C) | 2.4932 | 11.1127 |

TABEAU II.3 – Erreurs d’apprentissage et de validation pour des jeux expérimentaux avec un modèle système d’état d’ordre 5

En identification, on remarque une tendance globalement similaire, mais légèrement sur-estimé. Quant à la phase de validation, les premiers jours sont assez bien représentés mais assez rapidement, on remarque un décrochement et l’erreur peut atteindre jusqu’à 11 ° C à certains endroits.

Précédemment, la modélisation ARX avait donné de meilleurs résultats que la modélisation par système d’état. Voyons si c’est toujours le cas.

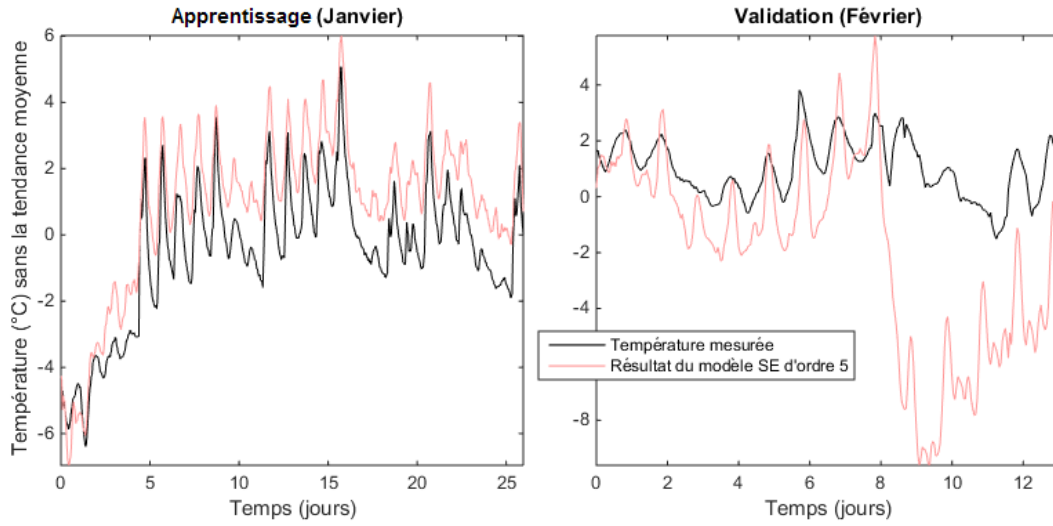


FIGURE II.17 – Prédiction sur un jeu de données expérimentales avec un système d'état d'ordre 5

- Identification avec un système ARX

Pour ces jeux de données, l'algorithme estime les ordres suivants :

$$\begin{aligned} n_a &= 1 \\ n_b &= [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1] \\ n_k &= [9 \ 9 \ 9 \ 9 \ 9 \ 9 \ 9 \ 9 \ 9 \ 9 \ 9 \ 9 \ 9] \end{aligned} \quad (\text{II.15})$$

Les résultats en prédiction sont observables sur la figure II.18. Le tableau II.4 donne les valeurs d'erreurs obtenues.

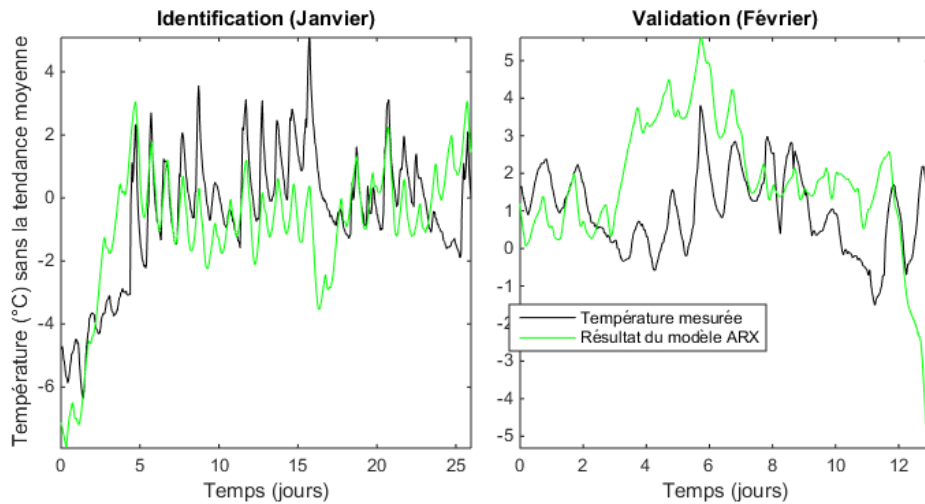


FIGURE II.18 – Prédiction sur un jeu de données expérimentales avec un modèle ARX

Cette fois-ci, ils sont moins bons que ceux obtenus avec le modèle système d'état.

Les difficultés d'identification rencontrées avec les entrées obtenues par mesures peuvent s'expliquer par la complexité des signaux qui est plus importante. De plus, des incertitudes

| | Identification (Janvier) | Validation (Février) |
|----------------|--------------------------|----------------------|
| Erreur moyenne | 1.4290 | 1.7922 |
| Erreur max | 4.6995 | 6.6133 |

TABLEAU II.4 – Erreurs d'apprentissage et de validation pour des jeux expérimentaux pour un modèle ARX

(entrées calculés approximativement à partir des signaux expérimentaux, erreur de capteurs, situations non attendus, etc.) peuvent apparaître et venir d'autant plus fausser les paramètres estimés pour le jeu d'identification.

En conclusion, les systèmes universels sont rapides à tester et à mettre en place. Cependant, il semble complexe de trouver un modèle unique qui convienne à toutes les situations. De plus, les résultats en prédiction, qui sont ceux auxquels nous nous intéressons tout particulièrement, sont souvent assez mauvais (plusieurs degrés d'erreur) ce qui peut s'avérer très gênant lorsque l'on veut faire de la gestion anticipative.

Enfin, nous allons voir dans la dernière partie que ces modèles se révèlent souvent instables.

II.3 Instabilité des modèles

Même dans le cas d'une bonne prédiction, comme nous avons pu en avoir avec un modèle ARX, le modèle n'est pas exploitable en gestion car instable. En discret, pour qu'un système soit stable, il faut que le module de tous les pôles soit strictement inférieur à 1. Or si on regarde la figure II.19, on voit ce que n'est pas le cas.

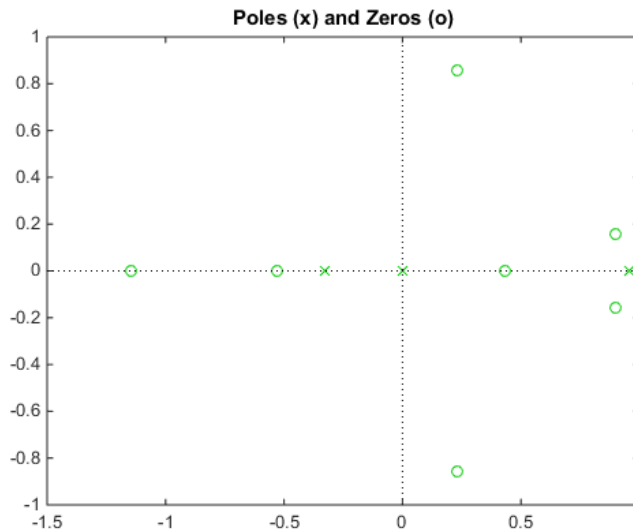


FIGURE II.19 – Pôles et zéros du modèle ARX

Les courbes de la figure II.20 montrent les résultats que nous pouvons obtenir lorsqu'il y a instabilité des pôles avec un modèle ARX.

Bien que des procédés permettent de garantir la stabilité d'un modèle ARX existent, ils sont souvent difficiles et contraignants à mettre en place [31].

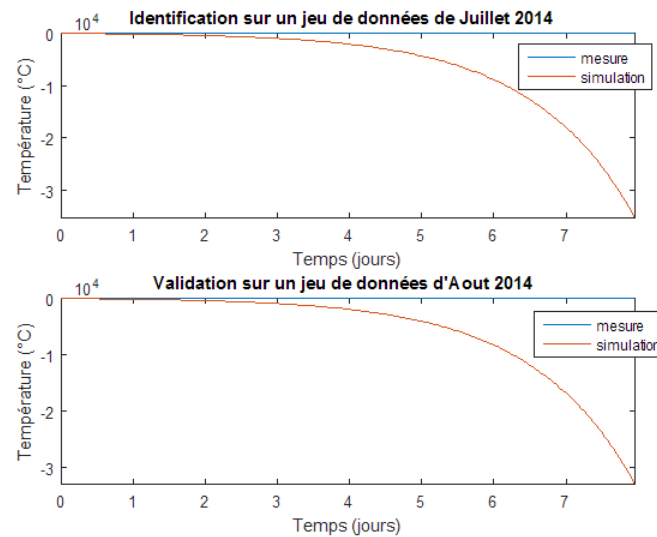


FIGURE II.20 – Identification et validation d'un modèle ARX non convergeant

II.4 Besoin de modèles non-linéaires

La plupart des processus que l'on peut rencontrer nécessitent des modèles non linéaires si l'on souhaite les décrire de manière précise dans la totalité de leur domaine de fonctionnement.

La présence d'une ventilation double flux pour renouveler l'air des salles PREDIS MHI induit des non linéarités. La plupart des modèles linéaires constituent des approximations valables dans un domaine plus ou moins restreint. Il est important de pouvoir élaborer un modèle non linéaire pour rendre compte du processus, non seulement autour de ses points de fonctionnement "habituels", mais également lors des passages d'un point de fonctionnement à un autre.

Ainsi, nous avons vu dans ce chapitre que la modélisation à l'aide de modèles universels présentait des limites. Dans le chapitre suivant, une modélisation à l'aide d'un modèle inspiré de la physique sera réalisée.

CHAPITRE III

Première approche de l'estimation paramétrique à base de modèle physique

Table des matières

| | |
|--|----|
| III.1 Avantages des approches à base de modèle physique | 55 |
| III.1.1 Intègre la connaissance disponible sur le système | 55 |
| III.1.2 Extrapolabilité et capacité à représenter des situations non représentées dans le jeu de données | 55 |
| III.1.3 Capacité du modèle à être validé par la physique | 56 |
| III.2 Description du modèle physique de PREDIS MHI | 56 |
| III.2.1 Confort des occupants | 57 |
| III.2.2 Modélisation thermique de PREDIS MHI | 60 |
| III.2.3 Les entrées du modèle : températures et puissances injectées | 64 |
| III.2.4 Les paramètres du modèle | 69 |
| III.3 Estimation paramétrique en utilisant un modèle physique selon une approche naïve | 71 |
| III.3.1 Simulation du modèle avec les valeurs initiales des paramètres | 71 |
| III.3.2 Première estimation des paramètres du modèle | 72 |
| III.4 Problématiques liées à l'estimation paramétrique en utilisant un modèle physique | 73 |
| III.4.1 Approche non-linéaire | 73 |
| III.4.2 Besoins de valeurs de paramètres initiales | 75 |
| III.4.3 Identifiabilité | 75 |
| III.4.4 Obtention d'un modèle physique adapté | 78 |

III.1 Avantages des approches à base de modèle physique

Avant même d'être identifié, un modèle construit à partir de la physique du processus étudié présente plusieurs avantages.

III.1.1 Intègre la connaissance disponible sur le système

La connaissance partielle d'un processus et de son environnement permet la réalisation d'un modèle lié à la physique. Les lois physiques représentent le processus dans un large ensemble de contextes. Elles se traduisent par des équations de modèle dont les paramètres représentent les caractéristiques intrinsèques du processus.

La mise en équation s'inspire de la réalité. La connaissance, même partielle, du processus aide à la réalisation du modèle, déterminant ainsi les variables d'influences et une structure de modèle. Comme pour un modèle universel, les paramètres peuvent être ajustés. Ces paramètres ont un sens physique, sont calculés en fonction de la connaissance du processus et recherchés dans un intervalle lié aux incertitudes et aux simplifications du modèle.

III.1.2 Extrapolabilité et capacité à représenter des situations non représentées dans le jeu de données

Lorsque l'on s'inspire de la physique pour concevoir un modèle, l'ensemble des contextes de fonctionnement y sont représentés. Cela traduit une capacité à représenter toutes sortes de situations, même celles qui ne sont pas présentes dans le jeu de données servant à l'identification.

Si on reprend les jeux de données simulées de Janvier et de Février issus du logiciel COMFIE-Pléiades utilisés dans le chapitre précédent et qu'on identifie les paramètres d'un modèle physique (qui sera décrit en détail dans la suite de ce chapitre), on remarque que lors du changement de scénario survenant dans le jeu de Février on obtient des résultats cohérents malgré le fait que cette situation (vacances scolaires où il n'y a ni occupation ni ventilation) ne soit pas représentée dans le jeu de données servant à l'identification (figure III.1).

| | Identification (Janvier) | Validation (Février) |
|----------------|--------------------------|----------------------|
| Erreur moyenne | 0.2279 | 0.4542 |
| Erreur max | 1.12 | 1.9925 |

TABLEAU III.1 – Erreurs d'apprentissage et de validation pour des jeux issus de simulation via COMFIE-Pléiades avec un modèle boîte grise

Bien qu'un peu moins précis, les résultats sont proches des meilleurs résultats que nous avons obtenus avec un modèle ARX, sans en présenter ses inconvénients énoncés au chapitre précédent.

Dans le cadre de la gestion anticipative, la connaissance des systèmes de ventilation, de chauffage et de refroidissement et de leur influence sur la température intérieure est primordiale. En effet, ce sont ces systèmes qui vont être pilotés pour réguler la température

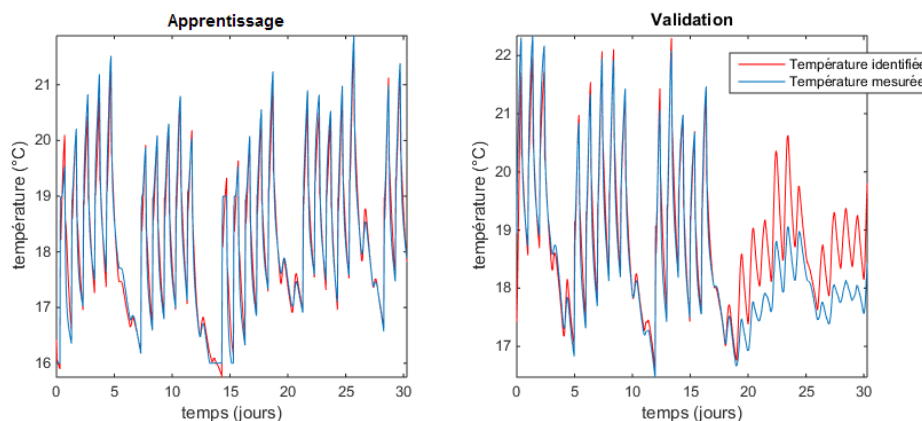


FIGURE III.1 – Identification et validation de jeux de données issus du logiciel COMFIE avec une modélisation physique

intérieure. En se focalisant sur la bonne représentation de ces systèmes dans la modélisation globale, on augmente la capacité du système à mieux représenter les situations absentes du jeu de données servant à l'identification.

De plus, l'espace de recherche est limité. En effet, des bornes acceptables par rapport au domaine de valeurs des paramètres sont déterminées et définissent le domaine de validité du modèle. Celui-ci est inspiré de la physique. Ces bornes participent à la cohérence du modèle physique car en empêchant les paramètres d'obtenir des valeurs absurdes par rapport à la réalité lors de l'identification du modèle, on garantit la bonne représentation du système dans son ensemble et non par rapport à un mode de fonctionnement particulier (lié au jeu de données d'apprentissage).

III.1.3 Capacité du modèle à être validé par la physique

Enfin, l'un des avantages des modèles inspirés de la réalité est que l'on peut reconnaître la physique du processus étudié. En effet, si l'on souhaite représenter un flux thermique traversant un mur, on pourra utiliser une résistance thermique dont la valeur dépend des matériaux qui constituent le mur. En utilisant des résistances en série, on peut représenter plusieurs couches de matériaux et en introduisant une capacité thermique, on introduit les notions d'inertie et de déphasage entre deux températures.

Les salles PREDIS MHI sont représentées à l'aide d'un circuit de résistances et de capacités thermiques. A travers ce schéma, on peut reconnaître la disposition des pièces et des zones adjacentes. Les flux de chaleur, les apports et les échanges entre les pièces sont donc aisément reconnaissables.

III.2 Description du modèle physique de PREDIS MHI

Dans la modélisation physique complète d'un bâtiment, on peut distinguer plusieurs aspects :

- la modélisation de l'enveloppe thermique du bâtiment ;

- la modélisation du comportement de l’occupant et la gestion de son confort ;
- la modélisation des équipements et systèmes énergétiques ;
- la modélisation de la qualité de l’air.

En premier lieu, il est intéressant d’étudier le confort des occupants. En effet, celui-ci définit la température de consigne idéale. Comment se traduit le confort thermique des occupants et quels sont les critères (saison, ensoleillement, activités) dont il dépend sont les questions que nous nous sommes posées pour le quantifier. Une fois les plages de température de confort identifiées, nous détaillerons notre modèle thermique dont la sortie est la température intérieure (température d’air) que nous essayons de réguler par la gestion anticipative. Les apports internes qui influent sur cette température ont également été étudiés. Les modèles d’équipements électriques simplifiés qui ont été utilisés afin d’obtenir les flux entrants de notre modèle seront ici détaillés.

III.2.1 Confort des occupants

Divers aspects peuvent influencer le confort des occupants d’une pièce. En voici une liste non exhaustive : la température, en premier lieu, mais également la luminosité, le taux de CO_2 , le taux d’humidité, le niveau sonore, la vitesse de déplacement de l’air (ventilateur), etc.

Ces facteurs, facilement quantifiables à l’aide de capteurs, n’assurent pas pour autant la satisfaction de l’occupant. En effet, le ressenti de l’occupant, et donc son confort, est propre à chacun et dépend de nombreux autres facteurs comme la saison, la météo, l’écart de température entre l’intérieur et l’extérieur, la présence d’une fenêtre dans la pièce, etc.

Les conditions physiologiques (température cutanée, sudation, vêtue, ...) nécessaires pour le confort thermique ont été déterminées par Fanger [14] sous des conditions thermiques homogènes et stationnaires. Si on écrit le bilan thermique, on peut exprimer la sensation thermique en fonction de l’écart du flux de chaleur cédé par le corps à l’environnement par rapport à celui qui correspond aux conditions de confort. Pour cela, on utilise l’indice PMV, qui exprime la sensation thermique moyenne éprouvée par un large groupe d’individus sur l’échelle de sensation thermique définie par l’ASHRAE [3] (cf. tableau III.2).

TABLEAU III.2 – Echelle de jugement ASHRAE

| |
|---|
| En ce moment précis, comment-ressentez vous l’air ambiant de ce local ? |
| 3 - froid |
| 2 - frais |
| 1 - légèrement frais |
| 0 - neutre |
| -1 - légèrement tiède |
| -2 - tiède |
| -3 - chaud |

La sensation thermique n’étant pas suffisante pour exprimer le confort, Fanger a proposé un autre indice qui complète le PMV, le « PPD » (pourcentage prévu des insatisfaits). Cet

indice permet de prévoir le pourcentage des insatisfaits à une sensation donnée. On peut exprimer le PPD en fonction du PMV, comme on peut le voir sur la figure III.2.

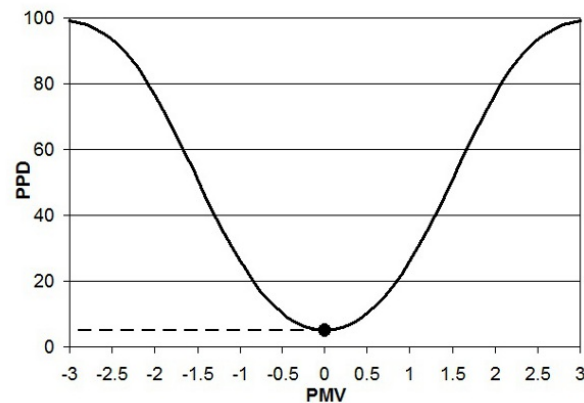


FIGURE III.2 – Pourcentage des insatisfaits (PPD) en fonction de la sensation thermique moyenne (PMV)

Afin de mieux cerner le confort des occupants de l'espace bureau de PREDIS, quelques statistiques ont été réalisées.

III.2.1-i Température de confort hiver

Nous avons demandé aux occupants de PREDIS MHI quelle était leur température idéale de confort en hiver, ainsi que les limites hautes et basses à ne pas dépasser selon eux. Ces différentes températures permettent de faire ressortir une tendance globale et de voir que tout le monde n'a pas les mêmes attentes que ce soit au niveau de la température de confort mais également au niveau de la plage de température qu'ils accepteraient de supporter. Tandis que les températures idéales sont relativement proches, les limites basses et hautes peuvent grandement varier.

Les figures III.3 montrent les résultats obtenus pour six occupants de l'espace bureau, en détail à gauche et en moyenne à droite. Les courbes ont été normalisées afin que toutes les aires soient égales.

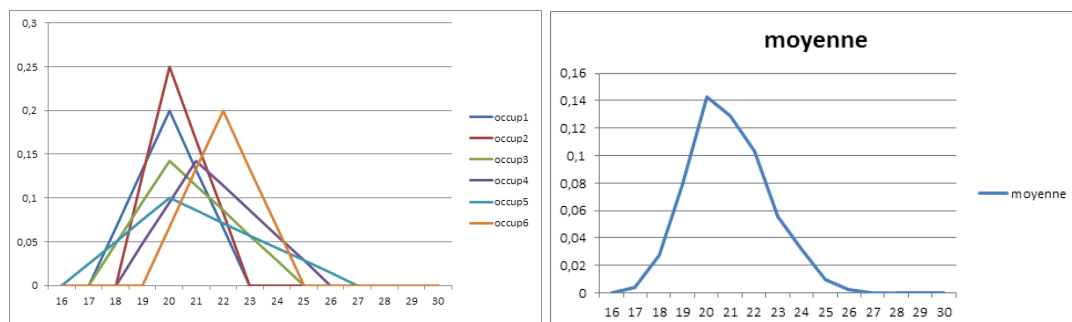


FIGURE III.3 – Plages de température acceptée par les occupants de l'espace bureau pendant l'hiver

Le maximum est atteint pour une température de 20°C, ce qui est une température de confort d'hiver relativement standard pour un espace bureau. En effet, la réglementation stipule une température de consigne de chauffage ne dépassant pas 19°C en hiver.

III.2.1-ii Température de confort été

Les températures de confort été et hiver sont généralement différentes. En effet, elle sont liées à la température extérieure, lorsqu'il fait plus froid, nous sommes mieux habillés et nous pouvons donc supporter une température intérieure plus basse, et inversement en été. Ainsi, la réglementation stipule une mise en marche des systèmes de climatisation lorsque la température intérieure dépasse les 26°C. En effet, chercher en été à atteindre les 19°C que l'on préconise pour l'hiver peut paraître aberrant lorsqu'il fait plus de 35°C dehors. La température de confort intérieure est donc grandement liée à la température extérieure et on peut trouver dans la littérature scientifique des modèles de température de confort variable qui dépendent de la température extérieure.

Voici en exemple l'un de ces modèles [4] :

$$T_{confort} = a * T_{ext} + b \quad (III.1)$$

avec $T_{confort}$ la température de confort (°C), T_{ext} la température extérieure (°C), a et b les paramètres du modèle à estimer.

Brager et de Dear [5] utilisent ce modèle avec les paramètres suivants (ce modèle de confort est valable pour une pièce non climatisée) :

$$T_{confort} = 0,31 * T_{ext} + 17,8 \quad (III.2)$$

Cette notion de confort adaptatif pourrait servir de température de consigne et permettre de limiter la consommation énergétique tout en préservant la santé des occupants. En effet, un écart de température trop grand avec l'extérieur (>10°C) en été amène des troubles respiratoires et l'utilisation excessive de la climatisation peut provoquer des allergies car l'air recyclé favorise la concentration des allergènes.

La figure III.4 compare la température extérieure à la température de confort obtenue avec la formule de l'équation III.2 pour un jeu de données collecté en Juin 2014.

La température de confort varie entre 22 et 28°C tandis que la température extérieure varie entre 15 et 40°C.

Maintenant que nous connaissons les températures moyennes de confort, nous cherchons à modéliser le bâtiment avec la température intérieure en sortie afin de pouvoir la réguler.

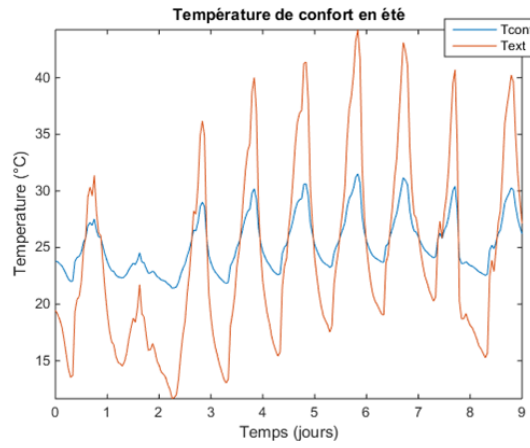


FIGURE III.4 – Température de confort Juin 2014

III.2.2 Modélisation thermique de PREDIS MHI

III.2.2-i Modélisation de l'enveloppe thermique

Les enveloppes thermiques sont représentées par des schémas de résistances et de capacités thermiques. Dans une telle modélisation, les résistances thermiques représentent l'isolation des murs tandis que les capacités thermiques figurent les effets dynamiques et donc l'inertie de la pièce.

Les phénomènes physiques qui interviennent dans cette représentation sont les suivants :

- La **conduction** : La conduction thermique est un transfert thermique spontané d'une région de température élevée vers une région de température plus basse. Elle est décrite par la loi de Fourier : la densité de flux de chaleur est proportionnelle au gradient de température :

$$\phi = -\lambda \text{grad}(T) \quad (\text{III.3})$$

la constante de proportionnalité λ est la conductivité thermique du matériau. Elle est toujours positive et s'exprime en $(W.m^{-1}.K^{-1})$. La densité de flux de chaleur ϕ s'exprime en $(W.m^{-2})$ et la température T s'exprime en kelvin (K).

Si on réalise un bilan d'énergie dans un corps homogène, l'expression de la loi de Fourier mène à l'équation générale de la conduction de la chaleur :

$$\lambda \Delta T + P = \rho c_p \frac{\partial T}{\partial t} \quad (\text{III.4})$$

où λ est la conductivité thermique du matériau $(W.m^{-1}.K^{-1})$, ΔT correspond au laplacien de la température, P est la densité volumique de puissance au sein du matériau $(W.m^{-3})$, ρ est la masse volumique du matériau $(kg.m^{-3})$ et c_p est la chaleur spécifique massique du matériau $(J.kg^{-1}.K^{-1})$.

Lorsque l'on souhaite calculer la résistance thermique équivalente d'une paroi située entre deux températures, trois phénomènes interviennent : la conduction, la convection et le rayonnement. La conduction est un phénomène qui intervient au sein du mur tandis que la convection et le rayonnement sont des phénomènes qui interviennent au

niveau des surfaces d'échanges. La résistante thermique d'une paroi peut s'exprimer comme ceci :

$$R_{th} = \frac{1}{h_{interface}^1} + \sum_{i=1}^N \frac{e_i}{\lambda_i} + \frac{1}{h_{interface}^2} \quad (III.5)$$

où N représente le nombre de couches de matériaux composant le mur, e_i correspond à l'épaisseur du matériau et λ_i sa conductivité thermique.

Les calculs pour obtenir les coefficients $h_{interface}$ sont détaillées en annexe A.

- Le **stockage** : En repartant de l'équation de la conservation de l'énergie pour une couche du mur de largeur dx et en appliquant la loi de Fourier, on arrive à une équation de la diffusion de la chaleur à une dimension :

$$\lambda \frac{\partial^2 T(t, x)}{\partial x^2} = \rho c_p \frac{\partial T(t, x)}{\partial t} \quad (III.6)$$

où T est la température de la couche qui dépend du temps et de la position selon x (épaisseur du mur).

On peut réaliser une approximation liée à la résolution temporelle et à l'horizon considéré pour obtenir les expressions des capacités thermiques des couches constituant un mur :

$$C = \frac{\rho c_p e S}{n} \quad (III.7)$$

avec n le nombre de couches constituant le mur.

- Le **transfert de masse** : ici le transfert de masse est modélisé par de la conduction à l'aide d'une résistance variable dont la valeur dépend du débit d'air de la VMC double flux.

Le modèle est discrétisé avec un pas de temps d'une heure. A cette échelle de temps, seules certaines dynamiques sont visibles. Le modèle de la salle de classe est d'ordre 1 et possède une seule capacité située au niveau de la dalle de béton. En effet, le béton est un matériau lourd et dense comparé aux murs et au plafond. Pour simplifier le schéma, toute l'inertie a donc été relocalisée au niveau de la dalle.

La figure III.5 représente le schéma qui a été retenu pour modéliser la salle de classe.

Par la suite, les calculs et matrices obtenues sont décrits pour la salle de classe. Les calculs et matrices de l'espace bureau seront obtenus par analogie.

Les équations obtenues aux nœuds A et B de la figure 5 donnent les relations suivantes :

$$\begin{aligned} \frac{dT_{dall}}{dt} &= \frac{1}{C} \left(\frac{T_{classe}}{R_1} + \frac{T_{EP}}{R_0} - T_{dall} \left(\frac{1}{R_0} + \frac{1}{R_1} \right) \right) \\ T_{classe} &= R_{equivalent} \left(\frac{T_{ext}}{R_{VMC}} + \frac{T_{shed}}{R_{shed}} + \frac{T_{off}}{R_{off}} + \dots \right. \\ &\quad \left. \frac{T_{cor}}{R_{cor}} + \frac{T_{EP}}{R_L} + \frac{T_{dall}}{R_1} + \phi_{classe} \right) \end{aligned} \quad (III.8)$$

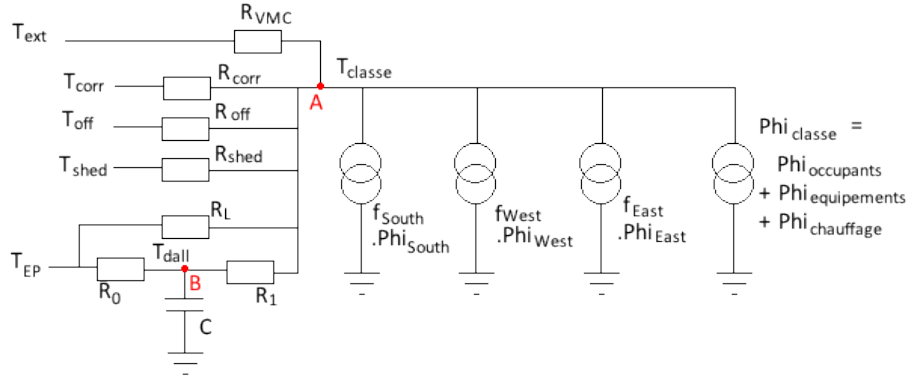


FIGURE III.5 – Schéma électrique équivalent de la salle de classe PREDIS MHI

avec

$$\phi_{classe} = f_{solar}^{East} \phi_{East} + f_{solar}^{South} \phi_{South} + f_{solar}^{West} \phi_{West} + \phi_{occupants} + \phi_{elec} + \phi_{chauffage}$$

$$R_{equivalent} = \frac{1}{\frac{1}{R_{VMC}} + \frac{1}{R_{shed}} + \frac{1}{R_{off}} + \frac{1}{R_{cor}} + \frac{1}{R_L} + \frac{1}{R_1}}$$

T_{corr} correspond à la température du couloir, T_{off} à la température de l'espace bureau, T_{shed} à la température du shed et T_{EP} correspond à la température de la zone de recherche du rez-de-chaussé.

En substituant l'équation de la température de la salle de classe de l'équation de la seconde ligne dans l'équation de la première on obtient l'équation différentielle suivante :

$$\begin{aligned} \frac{dT_{dall}}{dt} = & \frac{1}{C} \left(\frac{R_{equivalent}}{R_1} \left(\frac{T_{ext}}{R_{VMC}} + \frac{T_{shed}}{R_{shed}} + \frac{T_{off}}{R_{off}} + \frac{T_{cor}}{R_{cor}} + \Phi_{classe} \right) \right. \\ & \left. + \left(\frac{1}{R_1} + \frac{R_{equivalent}}{R_1 R_L} \right) T_{EP} - T_{dall} \left(\frac{1}{R_0} + \frac{1}{R_1} - \frac{R_{equivalent}}{R_1^2} \right) \right) \end{aligned} \quad (III.9)$$

En posant T_{dall} la variable d'état, $U = (T_{ext} \ T_{cor} \ T_{off} \ T_{shed} \ T_{EP} \ \phi_{classe})$, le vecteur des sollicitations et T_{classe} la variable de sortie alors on obtient le système d'état suivant :

$$\begin{aligned} \frac{dT_{dall}}{dt} &= A \times T_{dall} + B \times U(t) \\ T_{classe} &= C \times T_{dall} + D \times U(t) \\ t = 0, T_{dall}(0) &= T_{w0} \end{aligned} \quad (III.10)$$

avec

$$\begin{aligned}
 A &= \frac{1}{C} \times \left(-\frac{1}{R_0} + \frac{R_1 - R_{\text{équivalent}}}{R_1^2} \right) \\
 B &= \begin{pmatrix} \frac{R_{\text{équivalent}}}{R_1 \times R_{VMC}} \\ \frac{R_{\text{équivalent}}}{R_1 \times R_{cor}} \\ \frac{R_{\text{équivalent}}}{R_1 \times R_{off}} \\ \frac{R_{\text{équivalent}}}{R_1 \times R_{shed}} \\ \left(\frac{1}{R_0} + \frac{R_{\text{équivalent}}}{R_1 \times R_L} \right) \\ \frac{R_{\text{équivalent}}}{R_1} \end{pmatrix} \times \frac{1}{C} \\
 C &= \frac{R_{\text{équivalent}}}{R_1} \\
 D &= R_{\text{équivalent}} \times \begin{pmatrix} \frac{1}{R_{VMC}} \\ \frac{1}{R_{cor}} \\ \frac{1}{R_{off}} \\ \frac{1}{R_{shed}} \\ \frac{1}{R_L} \\ 1 \end{pmatrix}
 \end{aligned} \tag{III.11}$$

La VMC double flux est modélisée par une résistance variable qui dépend de l'efficacité de l'échangeur rotatif, de la quantité de fuite d'air et du débit d'air.

III.2.2-ii Modélisation de la VMC double flux

Une ventilation double flux assure le renouvellement de l'air, le chauffage en hiver et le rafraîchissement en été (il s'agit d'un free cooling). Il s'agit d'un système de ventilation équipé d'un moteur électrique qui assure automatiquement le renouvellement de l'air. L'extraction de l'air vicié et l'entrée d'air neuf sont contrôlés par deux ventilateurs. Cela permet d'intercaler entre ces deux flux un échangeur thermique, assurant la récupération de 60 à 65% des déperditions dues au renouvellement d'air [6]. La VMC double flux est constituée de deux parties :

- Une **partie électrique** et une **partie aéraulique** assurent le débit de l'air souhaité en fonctionnement normal et un débit maximal en cas d'incident.
- Une **partie thermique** qui assure la température souhaitée à l'intérieur du bâtiment.

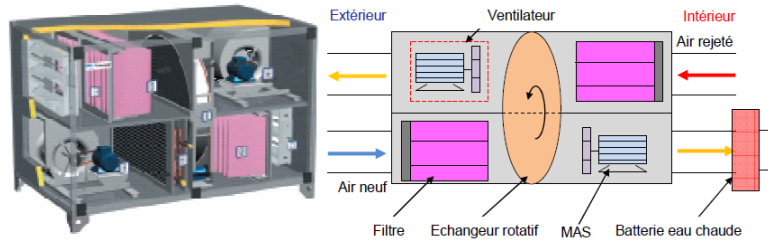


FIGURE III.6 – Schéma de la VMC double flux

Ces deux parties permettent de distinguer les variables contrôlables qui permettent de réguler la température des pièces : le débit d'air (quantité d'air) et le débit d'eau (température).

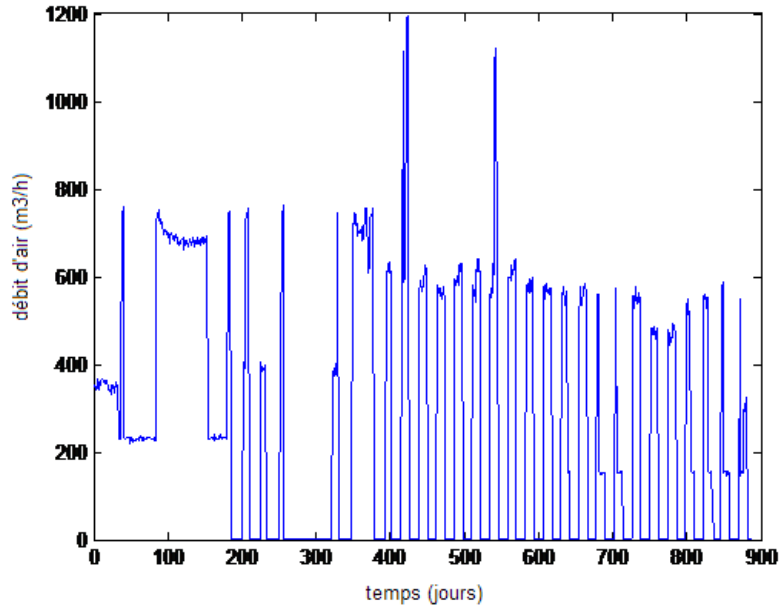


FIGURE III.7 – Débit d'air observé pour un jeu de donnée de 37 jours (Octobre/Novembre 2012)

Dans le modèle d'enveloppe thermique, la VMC est modélisée par une résistance variable fonction de la valeur du débit d'air dont l'équation est la suivante :

$$R_{VMC} = \frac{3600}{((1 - eff) \times C_{air} \times p_{air} \times (airLeak + Q_{airmode}))} \quad (III.12)$$

avec :

eff l'efficacité de l'échangeur rotatif,

C_{air} la chaleur massique de l'air,

p_{air} la densité de l'air,

$airLeak$ le coefficient de fuite d'air et

$Q_{airmode}$ le débit correspondant à l'un des quatre modes de ventilation définis.

III.2.3 Les entrées du modèle : températures et puissances injectées

On peut distinguer deux types d'entrées :

- Les températures : des zones adjacentes et extérieures, qui sont directement obtenues à l'aide des capteurs ;
- Les puissances injectées : qui sont obtenues par calcul ou interpolation et dont les calculs ont été décrits dans le chapitre II.

Bien que les calculs aient déjà été détaillés au chapitre précédent, nous allons revoir rapidement les modèles des équipements électriques.

III.2.3-i Les équipements électriques

Les appareils électriques des salles PREDIS sont les suivants :

- La VMC double flux et les batteries d'eau chaude qui renouvellent l'air et assurent un apport de chauffage en hiver ;
- Les batteries électriques des ordinateurs portables ;
- La pompe à chaleur de l'espace bureau qui permet de climatiser en été et de compléter le chauffage en hiver.

Ces dispositifs induisent des flux de chaleur qui interviennent en entrée du modèle de la salle de classe. Ces flux de chaleur sont les suivants :

- Le flux de chauffage ;
- Le flux provenant des équipements électriques ;
- Le flux lié à l'occupation et au métabolisme des corps ;
- Le flux solaire.

III.2.3.i - a) *Le flux de chauffage*

Comme vu précédemment, la VMC double flux est représentée par une résistance variable positionnée entre la source de température extérieure et la température intérieure que nous cherchons à modéliser. Ainsi, l'air entrant est dans un premier temps chauffé par l'air sortant à l'aide de l'échangeur rotatif. Ensuite, les batteries d'eau chaude complètent cet apport de chaleur afin de mettre l'air entrant à température ambiante.

La puissance de chauffage est obtenue par calcul à l'aide de la différence de températures amont et aval batterie multipliée par le débit d'eau dans la batterie, la capacité thermique et la masse volumique de l'eau, comme décrit dans le chapitre II, équation II.1.

Cette puissance est multipliée par un coefficient qui sera estimé à l'aide des mesures. Ce coefficient permet de mieux appréhender les erreurs de calcul.

III.2.3.i - b) *Les flux de chaleur issus des équipements électriques et de l'occupation*

Dans le chapitre précédent, les calculs permettant d'obtenir la puissance électrique sont détaillés. Ainsi, à partir de la consommation électrique liée à l'éclairage et aux ordinateurs, nous pouvons obtenir la puissance électrique qui intervient en entrée du modèle.

Nous avons vu que le calcul de cette puissance subissait de nombreuses approximations. Concernant le flux de personnes, la quantité de CO_2 n'étant pas connue, le flux électrique lié aux ordinateurs, qui est lié à l'occupation, peut être utilisé afin de quantifier ce flux.

Ainsi, on distinguera deux flux de chaleur :

- le flux provenant des équipements électriques qui n'est pas lié au nombre de personnes présentes : l'éclairage ;
- le flux provenant des équipements électriques qui est lié au nombre de personnes présentes : la consommation des ordinateurs portables.

Ces deux flux d'entrée seront tous deux multipliés par un coefficient multiplicateur qui sera estimé à l'aide des mesures afin d'ajuster leur valeur.

Le flux d'occupation sera ainsi lié à la consommation électrique, celui-ci sera donc multiplié par un coefficient à identifier qui permettra de représenter l'occupation. Celui-ci devrait donc logiquement être supérieur à 1 lors des périodes d'occupation.

III.2.3.i - c) Les flux solaires

Les flux solaires seront répartis en quatre flux :

- le flux solaire direct est ;
- le flux solaire direct sud ;
- le flux solaire direct ouest ;
- le flux solaire diffus.

Chacun de ces flux solaires est multiplié par un coefficient d'ajustement qui sera estimé à l'aide des mesures (les entrées du modèle). Ceci permet de rendre compte de la part de chaque flux solaire.

III.2.3.i - d) La pompe à chaleur

Le bureau est équipée d'une pompe à chaleur qui réchauffe l'atmosphère en hiver et la rafraîchit en été. La puissance électrique est directement mesurée à l'aide d'un capteur. De même que pour les autres flux, ce flux est multiplié par un coefficient estimé en fonction des mesures. Ce coefficient sera positif lorsque la pompe à chaleur est en mode chauffage et négatif lorsque la pompe à chaleur est en mode climatisation.

Dans le cadre de la gestion anticipative, nous cherchons à obtenir le meilleur compromis coût/confort. Ainsi, en période de chauffe, nous cherchons la température de consigne de chauffage qui permet d'atteindre le compromis recherché. Pour cela, dans un premier temps, les paramètres du modèle sont estimés à l'aide des données mesurées ou calculées. Ensuite, une prédiction de la température intérieure est réalisée avec les valeurs des paramètres estimées. Dans cette situation, les données d'entrée utilisées sont des données prédites.

La température extérieure et les flux solaires peuvent être anticipés à l'aide de prédictions météorologiques. Les flux internes peuvent également être anticipés car utilisés afin de réguler la température intérieure.

Ainsi, il reste les températures des zones adjacentes à anticiper. Pour cela, nous avons mis en place des modèles simplifiés faisant intervenir uniquement la température extérieure en entrée.

III.2.3-ii Les températures des zones adjacentes : prédiction des données d'entrées

Afin d'anticiper la température intérieure pour la mise en place de scénarios de gestion optimale, les entrées du modèle doivent également être prédites, mais à une échelle de précision moindre. La température extérieure et le flux solaire seront obtenus par prévision météorologique. Les apports internes peuvent également être anticipés (occupation) ou commandés pour la gestion (ventilation et chauffage).

Afin de prévoir l'évolution des températures des zones adjacentes, des modèles de prédiction simplifiés en fonction de la température extérieure ont été mis en place. Ceux ci ont été choisis de manière empirique pour retranscrire les variations des températures adjacentes en fonction de la température extérieure et de ses variations :

$$\frac{dT_{adjacente}}{dt} = a * \frac{dT_{ext}}{dt} + b * T_{ext} + c \quad (III.13)$$

Avec $T_{adjacente}$ la température de la zone à anticiper et T_{ext} la température extérieure. a , b et c sont les paramètres du modèle à identifier.

Pour chaque jeu de données, les coefficients a , b et c sont dans un premier temps identifiés sur les deux premiers tiers du jeu puis le modèle est validé sur le dernier tiers.

III.2.3.ii - a) Prédiction des températures de zones non occupées

Les zones non occupées sont le couloir et l'espace vide autour des salles PREDIS MHI. La température dans ces zones subit très peu les apports internes dus à l'occupation (chauffage et chaleur dégagée par le corps ou les ordinateurs). Ainsi, ces zones ont une courbe de température qui suit une tendance journalière proche de la température extérieure.

Pour deux jeux de données collectés à des périodes différentes (Janvier et Juin), la température du couloir a été estimée.

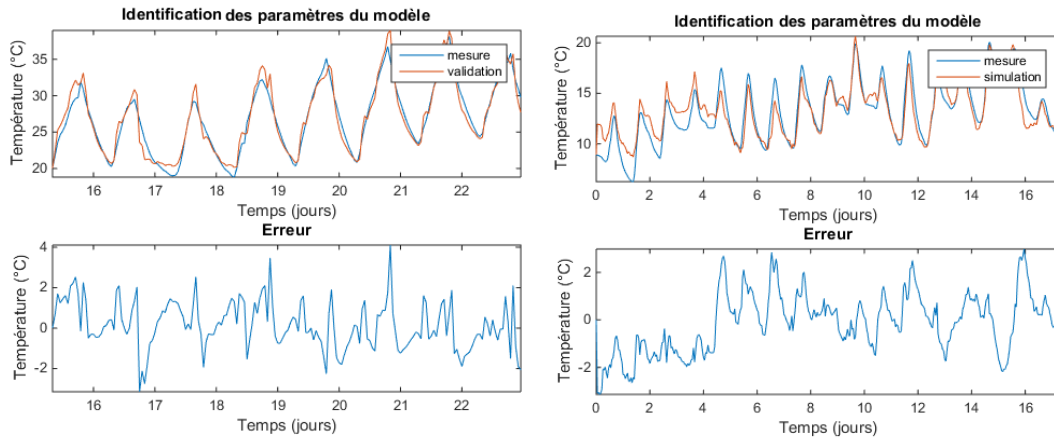


FIGURE III.8 – Identification de la température du couloir pour Juin et Janvier

| | a | b | c |
|---------|--------|--------|---------|
| Janvier | 0,1394 | 0,4442 | 10,0039 |
| Juin | 0,0305 | 0,6534 | 10,8755 |

TABEAU III.3 – Valeur des coefficients obtenus en identification pour des zones non chauffées pour deux jeux de données de saisons différentes

Comme on peut le voir sur les courbes des figures III.8 et III.9, les résultats sont très bons et suivent correctement les tendances en validation. Les valeurs obtenues pour les coefficients a , b et c sont visibles dans le tableau III.3. On remarque que les valeurs des paramètres b et c sont proches même si les jeux de données sont très différents (été et hiver). La valeur du paramètre a change en fonction de la saison.

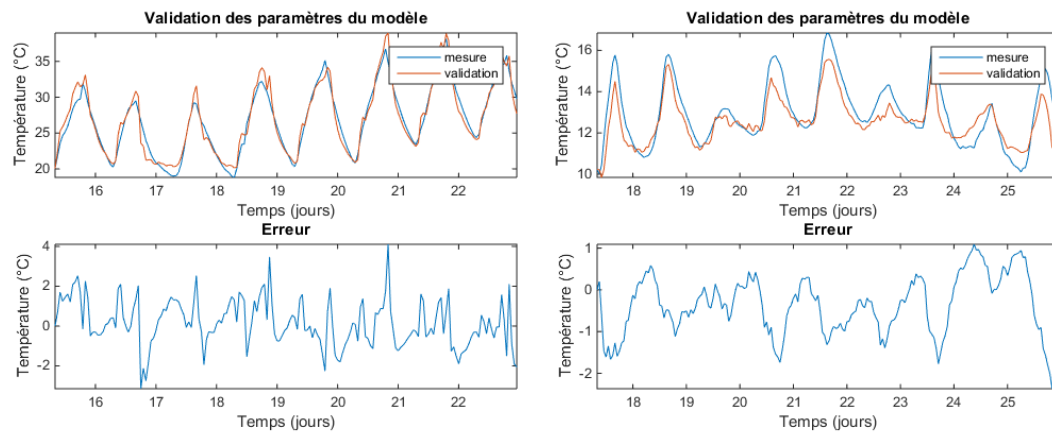


FIGURE III.9 – Validation de la température du couloir pour Juin et Janvier

Voyons les résultats pour des zones occupées.

III.2.3.ii - b) Prédiction des températures de zones occupées

Contrairement au couloir et à l'espace vide entourant les salles PREDIS MHI, la salle de recherche située au rez-de-chaussée sous les salles PREDIS MHI et les salles PREDIS MHI sont des zones occupées subissant les apports internes liés à l'occupation. Ainsi, ces zones ont une tendance journalière, liée à la température extérieure, mais également une tendance hebdomadaire du fait de l'absence d'occupation pendant le week end. Cette tendance peut être d'autant plus complexe lorsque la fréquentation des zones n'est pas régulière.

De même que précédemment, la température de la salle de recherche du rez-de-chaussée a été estimée pour deux jeux de données de saisons différentes (Janvier et Juin).

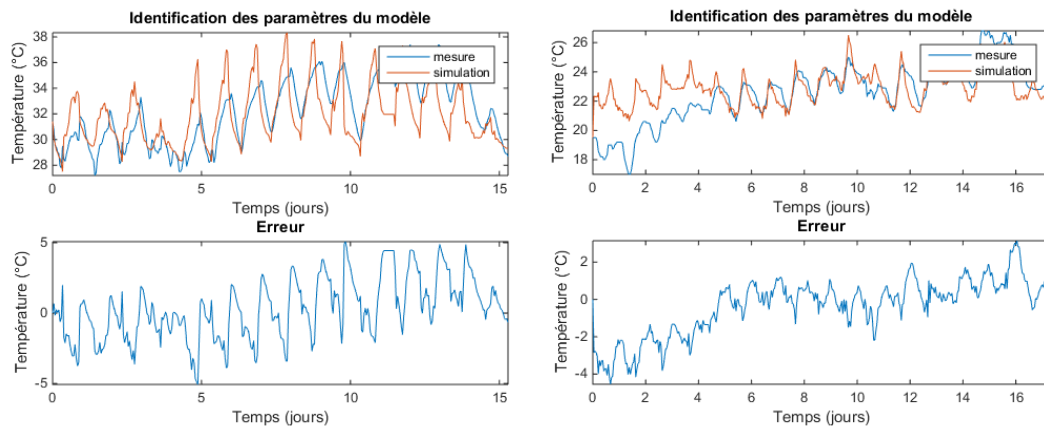


FIGURE III.10 – Identification de la température de la salle de recherche pour Juin et Janvier

| | a | b | c |
|---------|---------|--------|---------|
| Janvier | -0,135 | 0,3955 | 21,4379 |
| Juin | -0,1289 | 0,4193 | 25 |

TABLEAU III.4 – Valeurs des coefficients obtenus en identification pour des zones chauffées pour deux jeux de données de saisons différentes

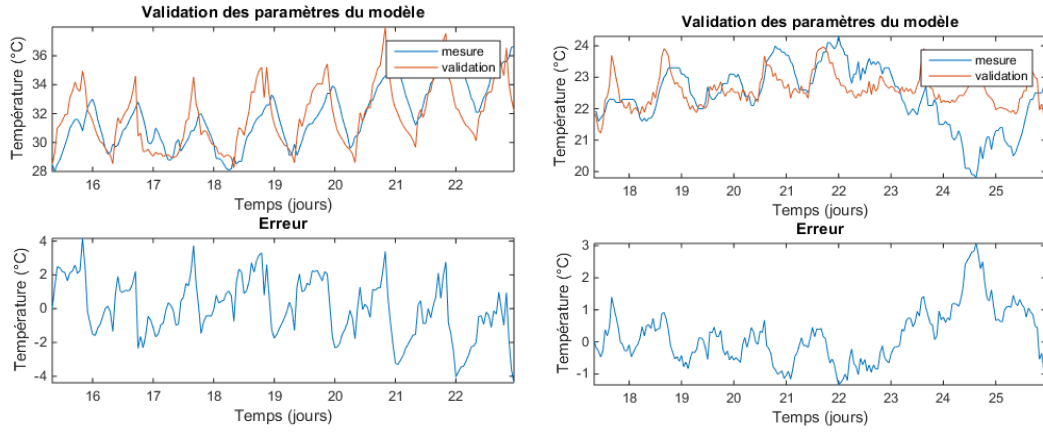


FIGURE III.11 – Validation de la température de la salle de recherche pour Juin et Janvier

De manière logique, les résultats, observables sur les figures III.10 et III.11 sont beaucoup moins bons. En effet, le comportement de ces pièces est beaucoup plus complexe à représenter avec pour seule entrée la température extérieure. Cependant, de même que précédemment, les valeurs des coefficients a , b et c sont proches quelle que soit la saison (tableau III.4).

A l'issue de ces expériences, on remarque donc qu'on pourrait mettre en place des modèles de température de zones adjacentes qui n'auraient pas nécessairement besoin d'être recalibrés en fonction de la saison. De plus, la représentation des zones chauffées pourrait être améliorée en prenant en compte une entrée en puissance qui modéliserait les apports internes attendus (occupation, périodes de chauffages, etc.).

III.2.4 Les paramètres du modèle

Le modèle initial de la salle de classe possède 19 paramètres :

— **les paramètres thermiques :**

- Résistances des murs entre la salle de classe et les zones adjacentes : $R_{corridor}$, $R_{offices}$, R_{space} , R_0 , R_1 et R_L ;
- Capacité thermique de la salle de classe (la dalle) : C ;

— **les coefficient liés aux apports internes :**

- coefficients de flux solaire : f_{solar}^{South} , f_{solar}^{East} , f_{solar}^{West} et f_{solar}^{Dif} ;
- $coef_{Occup}$: coefficient multiplicateur lié à l'apport de chaleur dû aux occupants ;
- $coef_{Chauff}$: coefficient multiplicateur lié à l'apport de chaleur dû au chauffage ;
- $coef_{Elec}$: coefficient multiplicateur lié à l'apport de chaleur dû au flux électrique ;

— **les coefficients liés à la ventilation double flux :**

- $efficiency$: efficacité de l'échangeur rotatif ;
- air_{Leak} : fuite d'air ;
- $coef_{Ventil}$: coefficient multiplicateur de la ventilation ;

— **La température initiale de la dalle : T_{w0} .**

Les coefficients multiplicateurs permettent d'ajuster les apports internes dont les calculs sont mal maîtrisés. Ces coefficients valent 1 si on utilise des signaux obtenus par simulations.

L'ensemble de ces paramètres constitue le vecteur des paramètres θ que nous cherchons à identifier.

Les équations du système d'état sont les suivantes :

$$\begin{aligned}\frac{dT_{dall}}{dt} &= A \times T_{dall} + B \times U(t) \\ T_{classe} &= C \times T_{dall} + D \times U(t) \\ t = 0, T_{dall}(0) &= T_{w0}\end{aligned}\tag{III.14}$$

III.2.4-i Calcul des valeurs nominales et définition des bornes d'incertitudes

Chacun de ces paramètres a été dans un premier temps calculé analytiquement, puis, nous avons défini des bornes d'incertitudes qui constituent des plages de validité pour chacun des paramètres. Ces plages de validité ont un sens physique par rapport à l'architecture du bâtiment et définissent l'espace de recherche des paramètres lors de l'identification.

| | valeur minimale | valeur initiale | valeur maximale |
|---------------------|--------------------|--------------------|--------------------|
| R_{space} | 0,012 | 0,017 | 0,028 |
| $R_{offices}$ | 0,078 | 0,13 | 0,21 |
| $R_{corridor}$ | 0,036 | 0,052 | 0,087 |
| R_L | 0,09 | 0,1007 | 0,11 |
| R_0 | 0,003 | 0,004 | 0,005 |
| R_1 | 0,0009 | 0,0013 | 0,002 |
| C | $1,32e^7$ | $2,01e^7$ | $2,89e^7$ |
| $efficiency$ | 0,5 | 0,8 | 0,85 |
| T_{w0} | 15 | 20 | 25 |
| $P_{residual}$ | 0 | 200 | 1000 |
| air_{Leak} | 10 | 22 | 30 |
| $coef_{occup}$ | 0,1 | 2 | 4 |
| $coef_{elec}$ | 0,1 | 1 | 3 |
| $coef_{chauff}$ | 0,1 | 1 | 3 |
| $coef_{ventil}$ | 0,1 | 1 | 2 |
| f_{solar}^{East} | 0 | 0,3 | 0,9 |
| f_{solar}^{South} | 0 | 0,3 | 0,9 |
| f_{solar}^{West} | 0 | 0,3 | 0,9 |
| f_{solar}^{Diff} | 0 | 0,3 | 0,9 |

TABLEAU III.5 – Valeur nominales, minimum et maximum obtenues par calcul et liées aux incertitudes

Les détails des calculs du tableau III.5 sont disponibles en annexe B.

III.2.4-ii Intérêt de la recalibration régulière d'un modèle

Lors de la création d'un modèle simplifié, de nombreuses approximations et simplifications sont réalisées. L'estimation des paramètres permet d'ajuster le modèle afin de le faire correspondre au mieux à la réalité. Cependant, une seule identification n'est pas suffisante pour

représenter le comportement global du bâtiment. En effet, chaque estimation paramétrique est réalisée à partir d'un jeu de données qui traduit des phénomènes bien précis.

Le réajustement des valeurs des paramètres permet ainsi une grande adaptabilité du modèle que ce soit par rapport aux phénomènes ponctuels ou annuels (saisons), mais également à l'évolution du bâtiment.

Cependant, l'estimation paramétrique nécessite une certaine expertise. En effet, une simple identification ne donne pas forcément des résultats satisfaisants du point de vue de la gestion anticipative.

III.3 Estimation paramétrique en utilisant un modèle physique selon une approche naïve

Avoir un modèle qui reflète la physique n'est pas suffisant pour garantir de bons résultats. En effet, ceci dépend de plusieurs facteurs : la complexité du modèle en cohérence avec la quantité d'informations contenue dans les jeux de données servant à l'identification mais également la manière dont est effectuée l'optimisation.

A l'aide de trois jeux de données issus de saisons différentes (un jeu de données d'hiver collecté en Janvier, un jeu de données mi-saison collecté en Avril et un jeu de données d'été collecté en Juillet), des tests en simulation et en identification ont été effectués. Le but de cette partie est de démontrer que les modèles physiques respectent une certaine représentation de la réalité mais que leur utilisation "naïve" (sans expertise) ne garantit pas la robustesse et la fiabilité du modèle.

Ici, la robustesse est la capacité de l'estimation paramétrique à retrouver le même résultat pour des points de départ différents. La fiabilité quant à elle est la capacité de l'estimation paramétrique à être validée avec un autre jeu de données.

III.3.1 Simulation du modèle avec les valeurs initiales des paramètres

Les paramètres sont calculés en fonction des caractéristiques du bâtiment. Certaines approximations ont été faites afin de simplifier le modèle.

Les figures [III.12](#) montrent les résultats lorsque les modèles sont simulés avec les valeurs initiales des paramètres.

Nous pouvons ainsi remarquer que la simulation respecte la même tendance que la mesure mais est souvent sur-estimée. Ceci peut être lié aux incertitudes sur les calculs des valeurs nominales des paramètres mais également aux simplifications et aux approximations de la structure du modèle. Cela peut induire un écart avec la réalité.

C'est pourquoi il est utile d'identifier les paramètres du modèle en cherchant à minimiser l'écart entre mesures et simulation. Des bornes de recherche sont définies afin de rester dans un domaine ayant un sens physique.

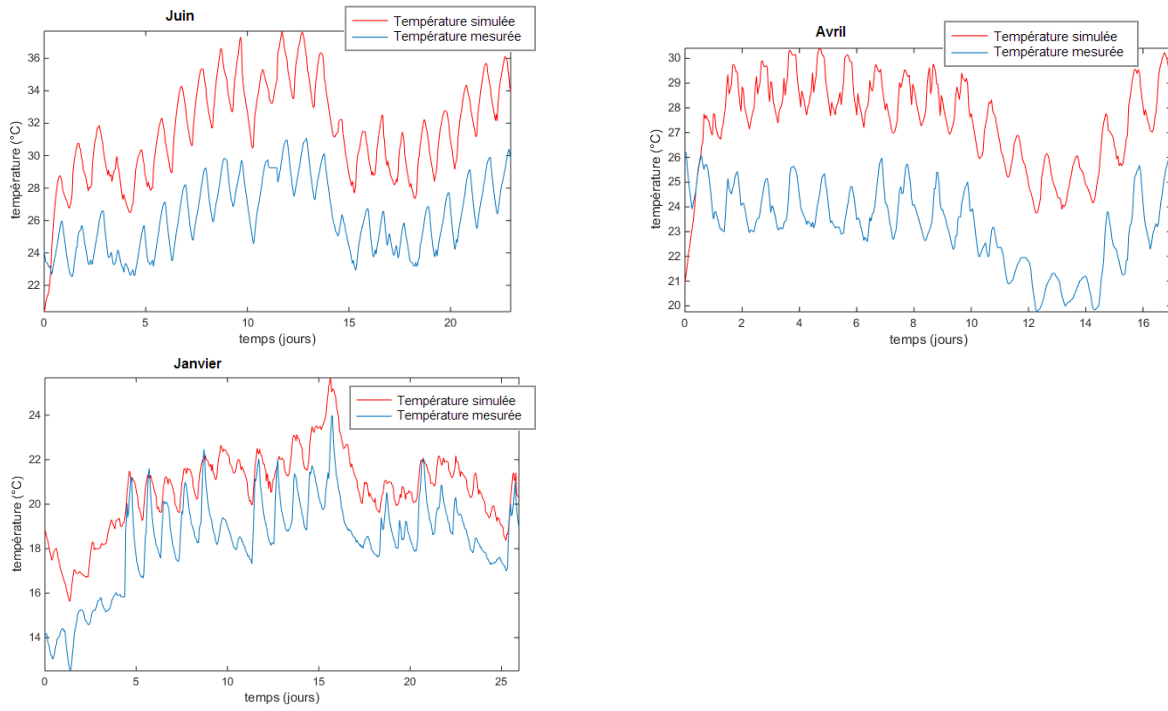


FIGURE III.12 – Simulation de la température intérieure de la salle de classe avec les valeurs initiales des paramètres pour des jeux de données de Janvier, Avril et Juin

III.3.2 Première estimation des paramètres du modèle

Pour plusieurs jeux de données, une identification a été réalisée afin de trouver un jeu de paramètres θ_i qui minimise l'écart entre température obtenue par simulation et température obtenue par mesure. Les paramètres initiaux sont obtenus par tirage aléatoire dans l'espace de recherche délimité par les bornes d'incertitudes.

$$erreur = \frac{\min}{\theta_i} \sum_{t=0}^{NT_s} |T_{mesure}(t) - T_{simulation}(\theta_i, U, t)| \quad (III.15)$$

avec N le nombre de points mesurés et

Il s'agit de minimiser l'erreur moyenne en valeur absolue entre la température obtenue par mesure (avec un pas d'échantillonnage $T_s = 3600s$) et la température obtenue par simulation.

Chacun de ces jeux de paramètres sera ensuite validé sur un autre jeu de données proche en comportement du jeu de données ayant servi pour l'identification.

Les figures III.13 montrent les résultats que nous obtenons respectivement en été et en hiver.

On remarque que les tendances sont globalement respectées mais que la validation pour le jeu d'été est bien meilleure que celle pour le jeu d'hiver. Ainsi, une simple estimation des paramètres n'est pas satisfaisante dans le cadre de la gestion anticipative. En effet, à certains endroits, l'erreur peut atteindre jusqu'à 3°C ou plus.

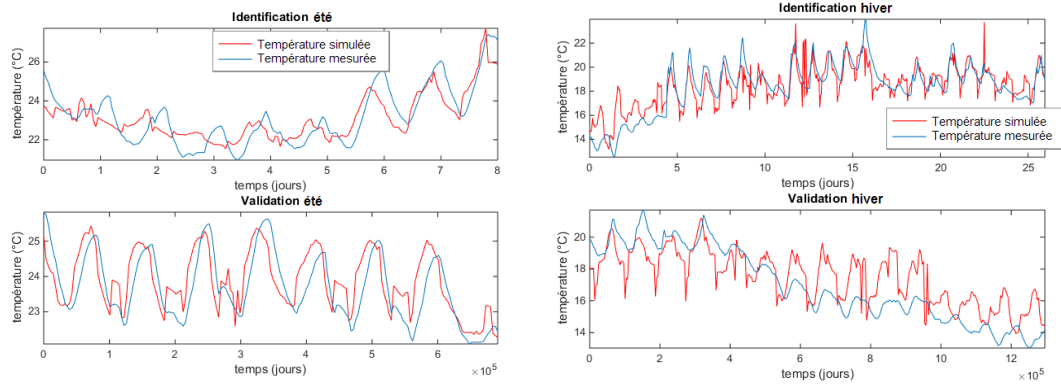


FIGURE III.13 – Identification et validation du modèle sur des jeux de données d'été et d'hiver

Nous allons donc observer quelles sont les problématiques liées à l'estimation paramétrique avec un modèle physique. Ceci nous permettra de mieux appréhender l'identification afin d'améliorer les résultats.

III.4 Problématiques liées à l'estimation paramétrique en utilisant un modèle physique

III.4.1 Approche non-linéaire

Comme énoncé dans le chapitre II, la présence d'une ventilation double flux pour renouveler l'air des salles PREDIS induit des non linéarités.

Généralement, on linéarise les modèles en se plaçant au niveau d'un point de fonctionnement. Cependant, ceci ne permet pas de modéliser un processus dans son ensemble, ni de représenter les passages d'un point de fonctionnement à un autre. En effet, lors de l'utilisation de modèles universels, le débit d'air de la ventilation était une entrée et les matrices du système d'état étaient constantes, le modèle était donc linéaire :

$$\dot{X}(t) = AX(t) + BU(t) \quad Y(t) = CX(t) + DU(t) \quad (\text{III.16})$$

Ici, le système d'état présenté dans la partie III.2.2-i de ce chapitre possède des matrices qui dépendent du temps car elles sont fonction du débit de la ventilation et des paramètres qui lui sont liés.

A chaque pas de temps t_k , on pourrait calculer la résistance variable représentant la ventilation en fonction de la valeur du débit d'air de la ventilation. Les matrices du système d'état seraient alors calculées au pas de temps t_k en fonction de la valeur de cette résistance.

$$\begin{aligned} A &= f_A(Q(t), efficiency, air_{Leak}, coef_{Ventil}) \\ B &= f_B(Q(t), efficiency, air_{Leak}, coef_{Ventil}) \\ C &= f_C(Q(t), efficiency, air_{Leak}, coef_{Ventil}) \\ D &= f_D(Q(t), efficiency, air_{Leak}, coef_{Ventil}) \end{aligned} \quad (III.17)$$

Nous avons des matrices qui changent à chaque pas de temps, le modèle est donc non linéaire. Ce genre de calcul se révèle extrêmement long si à chaque pas de temps il faut calculer la valeur de la résistance de la ventilation et les matrices du système d'état.

L'idée ici a été de simplifier la modélisation en définissant des modes de ventilation pour rendre le modèle linéaire par morceaux. Les valeurs des résistances de ventilation ainsi que les matrices du système d'état correspondantes ont été calculées. Ainsi, en fonction du mode correspondant à chaque pas de temps, les matrices sont appelées par l'algorithme.

Quatre modes de ventilation ont été définis :

- Un mode sans échange : $0 \text{ m}^3/h$;
- Un mode avec peu d'échange : $300\text{m}^3/h$ pour la salle de classe et $250\text{m}^3/h$ pour le bureau ;
- Un mode avec beaucoup d'échange : $600\text{m}^3/h$ pour la salle de classe et $450\text{m}^3/h$ pour le bureau ;
- Un mode "free cooling" pour un niveau de ventilation maximum : $1200\text{m}^3/h$ pour la salle de classe et $750\text{m}^3/h$ pour le bureau.

A chaque pas de temps, le mode de ventilation est déterminé en fonction de la valeur de débit d'air dont il est le plus proche.

Sur la figure III.14, on peut observer la différence entre une ventilation par paliers répartie sur les quatre modes définis en fonction de la valeur de la ventilation au pas de temps t et la ventilation réelle.

Dans le cas présent, le fait d'utiliser des paliers pousse à approximer le débit de ventilation à $300\text{m}^3/h$ alors que la mesure est aux alentours de $170\text{m}^3/h$. Cependant, si on compare les résultats de la figure III.15 où on a réalisé deux tests d'identification, un avec paliers et un sans, on remarque que les résultats sont quasiment les mêmes. Par contre l'identification sans définition des modes de ventilation prend environ 1350 secondes tandis que l'identification avec les modes de ventilations prend environ 109 secondes. Le temps de calcul est divisé par 13.

Les valeurs des paramètres estimées sont également relativement proche (cf. tableau III.6). On remarque quelques écarts qui sont probablement dus à la dispersion paramétrique et qui dépendent donc du point de départ qui a été tiré aléatoirement. On note malgré tout que les

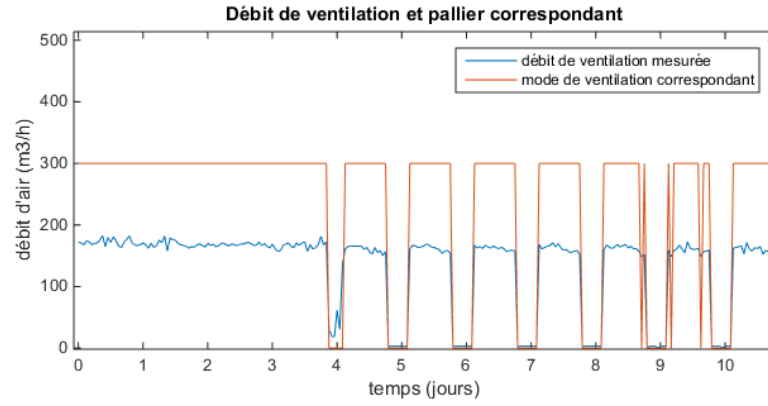


FIGURE III.14 – Comparaison de la mesure du débit d'air avec les paliers retenus pour un jeu de Janvier pour la salle de classe

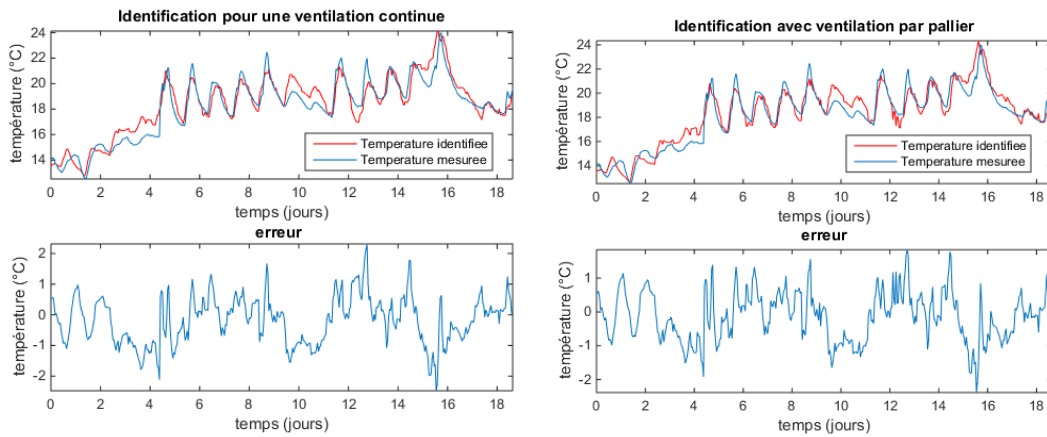


FIGURE III.15 – Identification des paramètres pour un modèle avec et sans discrétisation par paliers de la ventilation

paramètres impactés par ces petits écarts sont les paramètres liés à la ventilation : air_{Leak} , $efficiency$ et $Coeff_{ventil}$.

III.4.2 Besoins de valeurs de paramètres initiales

Enfin, contrairement aux modèles universels, les modèles physiques nécessitent des valeurs initiales de paramètres.

Ceci peut être problématique car cela implique des calculs qui demandent une connaissance assez précise du bâtiment, de ses dimensions et des matériaux qui le composent. L'utilisation de calculs analytiques et la définition de bornes de recherche découlant des incertitudes liées aux calculs analytiques garantissent un espace de recherche valide compte tenu de la physique du bâtiment.

III.4.3 Identifiabilité

À l'issue d'une série d'identifications avec différentes valeurs initiales de paramètres tirées aléatoirement dans l'espace de recherche, des statistiques peuvent être réalisées. En effet, les

| | Valeur minimum | Identification sans modes de ventilation | Identification avec modes de ventilation | Valeur maximum |
|---------------------|----------------|--|--|----------------|
| R_{space} | 0,0119 | 0,0119 | 0,0119 | 0,0284 |
| $R_{offices}$ | 0,0779 | 0,0779 | 0,0779 | 0,206 |
| $R_{corridor}$ | 0,0358 | 0,0358 | 0,0358 | 0,0874 |
| R_L | 0,09 | 0,11 | 0,11 | 0,11 |
| R_0 | 0,003 | 0,005 | 0,005 | 0,005 |
| R_1 | 0,0009 | 0,002 | 0,002 | 0,002 |
| C | 1,32e7 | 2,06e7 | 1,78e7 | 2,98e7 |
| $efficiency$ | 0,5 | 0,8496 | 0,5536 | 0,85 |
| f_{solar}^{East} | 0 | 0 | 0 | 0,9 |
| f_{solar}^{South} | 0 | 0 | 0 | 0,9 |
| f_{solar}^{West} | 0 | 0 | 0 | 0,9 |
| f_{solar}^{Diff} | 0 | 0 | 0 | 0,9 |
| T_{w0} | 15 | 15 | 15 | 35 |
| $P_{residual}$ | 0 | 84,072 | 0 | 1000 |
| air_{Leak} | 10 | 18,5281 | 30 | 30 |
| Coe_{fOccup} | 0,1 | 2,2984 | 2,5526 | 2 |
| Coe_{fElec} | 0,1 | 0,5953 | 0,2506 | 2 |
| $Coe_{fChauff}$ | 0,1 | 0,1 | 0,1 | 200 |
| $Coe_{fVentil}$ | 0,1 | 0,4725 | 1 | 2 |

TABLEAU III.6 – Valeurs des paramètres estimés dans l’espace de recherche pour les deux modèles (avec ou sans définition des modes de ventilation)

valeurs d’erreur moyenne peuvent être observées ainsi que les valeurs de paramètres identifiés obtenues.

On peut observer la dispersion obtenue pour 10 identifications avec les jeux d’hiver et d’été sur les figures III.16 et III.17. Il s’agit de toutes les valeurs moyennes d’erreur obtenues pour les 10 identifications. L’occurrence des erreurs est observable en ordonnée.

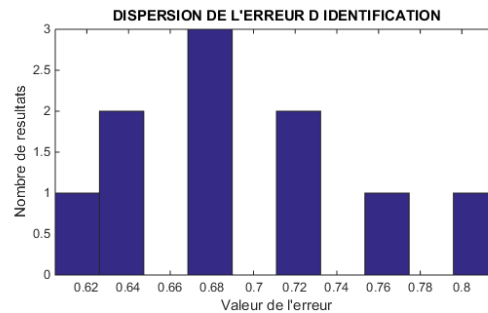


FIGURE III.16 – Histogramme des erreurs obtenues pour 10 identifications avec 10 points de départ différents pour un jeu d’été

Bien que les résultats soient globalement bons, l’erreur est dispersée. Cela traduit un manque d’ergodicité dans l’identification. En effet, pour des valeurs initiales différentes, l’algorithme ne converge pas vers la même solution.

On arrive aux mêmes conclusions par rapport aux valeurs des paramètres estimés, elles sont dispersées.

Les histogrammes de la figure III.18 correspondent à la dispersion paramétrique de chaque paramètres pour 10 identifications réalisées. En abscisse, on observe la valeur estimée tandis

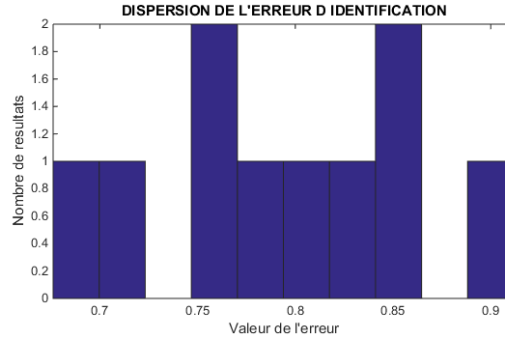


FIGURE III.17 – Histogramme des erreurs obtenues pour 10 identifications avec 10 points de départ différents pour un jeu d'hiver

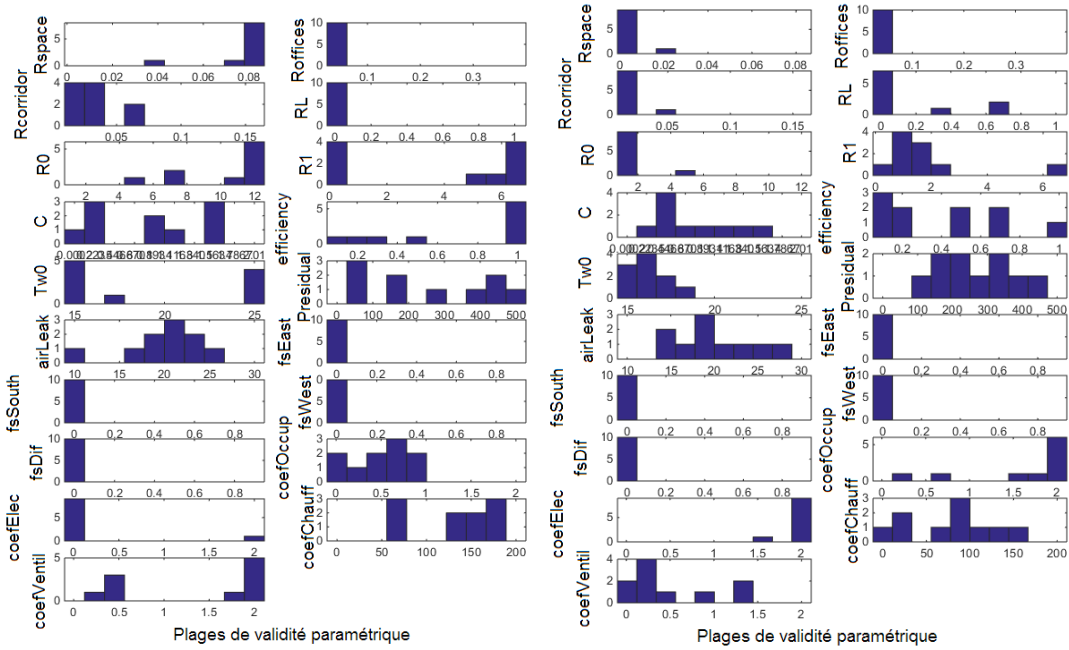


FIGURE III.18 – Dispersion paramétrique pour 10 identifications avec 10 points de départ différents pour un jeu d'hiver et un jeu d'été

que l'ordonnée correspond au nombre d'occurrences de cette valeur sur les 10 identifications. On peut observer que certains paramètres sont très dispersés. Parmi les plus dispersés, on retrouve les paramètres :

- C : la capacité thermique de la dalle de béton ;
- $P_{residual}$: la puissance résiduelle ;
- air_{Leak} : le coefficient de fuite d'air de la VMC double flux ;
- $coef_{occup}$: le coefficient multiplicateur lié au flux de personnes.

Ces paramètres sont liés à des phénomènes qu'il est difficile de modéliser, d'autant plus que le niveau de finesse de notre modèle ne permet pas de bien les prendre en compte. Ceci traduit un problème d'identifiabilité de certains paramètres.

Si on regarde les résultats en identification et en validation (figures III.19), où l'on a gardé le jeu de paramètres qui, parmi les 10 optimisations réalisées, minimisait l'écart entre

mesure et simulation, on remarque une amélioration par rapport aux résultats obtenus avec une approche naïve.

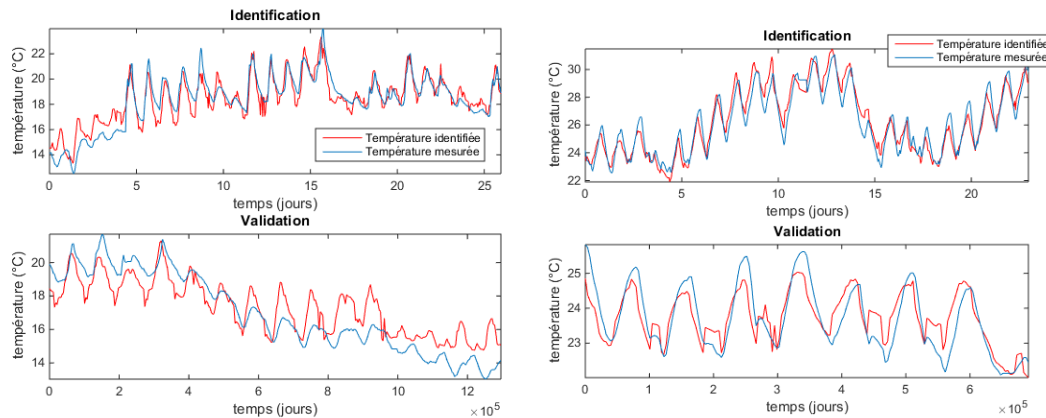


FIGURE III.19 – Meilleure identification parmi 10 et validation du modèle sur des jeux de données d’hiver et d’été

Ceci est logique puisque sur une série d’identifications, on garde le jeu de paramètres qui donne le meilleur résultat. Cependant, ce protocole d’identification est relativement long puisqu’il implique un certain nombre d’optimisations.

III.4.4 Obtention d’un modèle physique adapté

Le choix du modèle et les simplifications dépendent de l’application voulue. Ici, nous souhaitons établir des modèles pour la gestion anticipative à court terme (24, 48 ou 72h). Un modèle discret à pas de temps horaire a été retenu. Il existe un optimum entre le nombre de paramètres (ie. la complexité du modèle) et la précision en validation.

Lorsqu’un modèle est très détaillé, il donnera de bons résultats en identification, mais il est probable qu’il donne de moins bons résultats en validation car avec beaucoup de degrés de liberté, il va probablement sortir de son domaine de validité physique pour coller aux mesures.

D’un autre côté, la richesse des sollicitations doit être suffisante pour pouvoir solliciter tous les modes de fonctionnement présents dans le modèle. Si ces derniers contiennent peu d’informations (ceci peut être dû à une faible excitation lors de l’enregistrement des signaux), il sera difficile de savoir si les résultats d’identification reflètent les caractéristiques du processus ou bien des phénomènes que nous ne maîtrisons pas (erreur de capteurs, situations particulières, ouverture des portes, etc.)

La complexité choisie pour le modèle doit également être en accord avec l’application recherchée.

CHAPITRE IV

Analyse des problématiques liées à l'estimation paramétrique à base de modèle physique

Table des matières

| | |
|---|-----|
| IV.1 Estimation des paramètres | 81 |
| IV.1.1 Identifiabilité des paramètres | 82 |
| IV.1.2 Analyse de sensibilité | 83 |
| IV.1.3 Analyse de la dispersion paramétrique et de l'arrivée en butée des valeurs identifiées | 88 |
| IV.2 Étude de la richesse spectrale des signaux d'entrée | 89 |
| IV.2.1 Sollicitation de la pompe à chaleur de l'espace bureau à l'aide d'une séquence binaire pseudo-aléatoire | 91 |
| IV.2.2 Longueur des données nécessaires à l'identification | 95 |
| IV.2.3 Pertinence du choix des bornes sur les résultats d'estimation paramétrique | 98 |
| IV.3 Incertitudes sur les phénomènes | 99 |
| IV.3.1 Influence de l'état d'ouverture des portes sur l'identification de la température intérieure de PREDIS MHI | 100 |
| IV.3.2 Influence des scénarios de ventilation sur l'identification des salles PREDIS MHI | 101 |
| IV.4 Besoin de "méta-optimisation" | 102 |
| IV.4.1 Amélioration de l'identifiabilité paramétrique | 102 |
| IV.4.2 Etude des phénomènes mal maîtrisés | 103 |

Lorsqu'on réalise une estimation paramétrique, plusieurs aspects peuvent venir perturber l'identification.

Dans un premier temps, on peut s'intéresser à l'identifiabilité des paramètres, c'est à dire la capacité des paramètres à être identifiés ou non. Cette notion d'identifiabilité est souvent liée aux analyses de sensibilité paramétrique [10]. L'étude de ces aspects est capitale dans un problème d'identification. En effet, nous avons vu précédemment que la complexité du modèle était liée à sa structure et donc au nombre de paramètres à identifier ; et influait sur la qualité de l'identification. En pratique, le manque d'identifiabilité du système lors d'une série d'estimations paramétriques partant de valeurs initiales différentes est traduit par :

- La dispersion du critère à minimiser à l'issue d'une série d'identifications (présence de minima locaux) ;
- la dispersion des valeurs des paramètres identifiés sur l'espace de recherche ;
- la sortie du domaine de validité des valeurs identifiées des paramètres.

Ce manque d'identifiabilité peut être lié à la complexité de la structure mais également aux informations contenues dans les signaux d'entrée. Si les signaux d'excitation sont de faible amplitude, il en résulte des mesures fortement entachées par le bruit de mesure. Généralement, on s'intéresse au contenu spectral des signaux d'entrée. La séquence binaire pseudo aléatoire qui est constituée d'une succession de paliers modulés en largeur est le signal d'excitation le plus couramment utilisé. En effet, il se rapproche d'un bruit blanc et a donc un contenu riche en fréquences. Ce genre d'expérimentation peut améliorer ponctuellement l'identification et nous permettre de mieux comprendre le système. C'est pourquoi elle a été menée sur l'espace bureau à l'aide du pilotage de la pompe à chaleur en mode climatisation sur une période estivale.

Enfin, certains phénomènes non mesurés (ouverture des portes par exemple) peuvent influencer la qualité de l'identification. Ainsi, nous nous intéresserons à l'impact de diverses situations : changement dans un scénario de ventilation, ouverture des portes, etc. et nous essayerons de quantifier leur influences sur la variable de sortie.

IV.1 Estimation des paramètres

Lors de l'identification des systèmes, on se pose la question de l'identifiabilité des paramètres. En définissant la notion d'identifiabilité paramétrique, nous observerons qu'elle est étroitement liée à la sensibilité des paramètres et à leur impact sur la variable de sortie. De plus, nous observerons comment se traduit en pratique un manque d'identifiabilité.

IV.1.1 Identifiabilité des paramètres

IV.1.1-i Notion d'identifiabilité

Il existe de nombreuses définitions de l'identifiabilité des paramètres dans la littérature scientifique. On peut distinguer quatre catégories en fonction des conditions d'étude [10]. Généralement, il s'agit de vérifier que deux vecteurs de paramètres différents n'entraînent jamais le même comportement entrée-sortie. Ainsi, on dira qu'un paramètre θ_i est identifiable en $\theta^* \in \mathbb{P}$ si $\exists(u, x_0) \in \mathbb{U} \times \mathbb{R}^{n_x}$ tel que :

$$\forall \theta \in \mathbb{P} \text{ et } \forall t \in \mathbb{T}, y(t, u, x_0, \theta) = y(t, u, x_0, \theta^*) \Rightarrow \theta = \theta^* \quad (\text{IV.1})$$

Où \mathbb{P} correspond à l'espace de recherche des paramètres, \mathbb{U} à l'espace des sollicitations, \mathbb{T} au temps et \mathbb{R}^{n_x} à l'espace du vecteur d'état. y est le vecteur des sorties observées.

Lorsque cette vérification est effectuée avant toute expérimentation, dans un cadre idéal, on parle d'**identifiabilité a priori** (le système et le modèle ont des structures identiques, les signaux d'entrée-sortie sont non bruités et l'entrée et les instants de mesures sont indépendants des paramètres et des autres variables). Ce cadre théorique idéal a été défini par Walter et Pronzato [36]. Une fois l'identifiabilité a priori vérifiée, on se place alors dans un cadre expérimental (les signaux d'entrée-sortie sont non bruités et le signal d'entrée et les instants de mesures sont fixés) : on parle d'**identifiabilité a posteriori**. Enfin, si l'on prend en compte le bruit sur les données, nous parlerons d'**identifiabilité pratique** [11]. La dernière classe d'identifiabilité concerne l'**identifiabilité en probabilité**, où est étudiée la convergence en probabilité de l'estimateur des paramètres.

IV.1.1-ii Méthode de vérification d'identifiabilité a posteriori

Il existe de nombreuses méthodes d'étude de l'identifiabilité d'un modèle. Nous nous intéresserons particulièrement au développement en séries de Taylor, qui propose une vérification de l'identifiabilité en lien avec les fonctions de sensibilité des paramètres.

Reprenons la structure de modèle définie en chapitre I (I.1). Pour simplifier les calculs et les notations, nous supposons qu'il possède une entrée et une sortie et qu'il est a priori identifiable. Notons θ_0 le vecteur des paramètres du modèle nominal. Supposons que le modèle possède n_p paramètres à identifier. Si nous prenons un vecteur θ au voisinage du vecteur des valeurs nominales et que nous notons l'écart entre ces deux vecteurs $d\theta = \theta - \theta_0$ alors le développement en série de Taylor au premier ordre sur la sortie donne :

$$\forall t_k, y(t_k, \theta_0 - d\theta, u) = y(t_k, \theta_0, u) + \sum_{i=1}^{n_p} \frac{\partial y(t_k, \theta, u)}{\partial \theta_i} \Big|_{\theta_0} d\theta_i + O(d\theta^2) \quad (\text{IV.2})$$

D'après la condition de continuité d'un problème inverse, une variation locale de la valeur des paramètres est censée induire une variation de la sortie y . Or, lorsque l'on réalise une analyse de sensibilité paramétrique, c'est justement l'impact d'une variation de la valeur d'un paramètre sur la variable de sortie que l'on souhaite observer et quantifier.

IV.1.2 Analyse de sensibilité

IV.1.2-i Classes d'analyse de sensibilité

L'étude de la sensibilité d'un modèle vise à quantifier l'influence d'un facteur sur une réponse. Ici, les facteurs seront les paramètres et la réponse, la variable de sortie du modèle. Il existe trois classes d'analyse de sensibilité, selon [33] : les méthodes locales (autour d'un point nominal des paramètres), les méthodes de criblage et les méthodes globales (tout l'espace paramétrique est considéré).

Les **analyses de sensibilité locales** consistent en l'observation de l'influence d'une variation d'un paramètre autour d'une valeur nominale θ_0 sur la variable de sortie. Les autres paramètres sont fixés à leur valeur nominale. Ici, les interactions entre les paramètres ne sont pas prises en compte.

Les **méthodes globales** sont des méthodes probabilistes basées sur la connaissance de la distribution des paramètres sur tout l'espace. Elles permettent de classer les paramètres selon leur influence sur la variable de sortie tout en prenant en compte les interactions. Par exemple, les indices de Sobol mesurent la contribution d'un paramètre ou d'un groupe de paramètres dans un modèle. Cet indice est mesuré en divisant la variance de l'espérance conditionnelle du modèle lors de la variation du groupe de paramètres sur la variance "globale" du modèle.

Les **méthodes de criblages** permettent de repérer parmi un ensemble de paramètres potentiellement influents, ceux qui le sont dans un domaine de variation fixé. Ici, les paramètres n'ont pas de lois de distribution. Nous nous intéresserons à la méthode de MORRIS. La méthode de MORRIS [29] est une méthode qualitative qui permet de classer les paramètres en trois groupes en fonction de leur effets :

- Les effets négligeables ;
- Les effets linéaires sans interaction ;
- Les effets non-linéaires avec ou sans interaction.

La méthode de MORRIS mesure la sensibilité de chaque paramètre à l'aide de l'espérance des variations. Si la valeur est importante alors les effets sont en moyenne importants et la sortie du modèle est donc sensible aux variations. Cette méthode mesure également les interactions et les effets non-linéaires à l'aide de l'écart type. Si la valeur est grande, alors les effets sont différents les uns des autres et peuvent dépendre des entrées (effets non-linéaires) ou bien des autres paramètres (interactions).

IV.1.2-ii Application sur PREDIS MHI

IV.1.2.ii - a) Méthode locale

Comme décrit précédemment, cette méthode consiste en l'observation de l'influence des variations de paramètres sur la variable de sortie tandis que tous les autres sont respectivement fixés à leurs valeurs nominales.

Pour chaque paramètre, l'intervalle de recherche est discrétisé en N valeurs. Pour chacune de ces valeurs, le modèle est évalué, tandis que les autres valeurs de paramètres sont nominales. Soit l'équation suivante :

$$T_{sensible}^{(i,j)} = modele(\theta^{(i,j)}, U(t), t) \quad (IV.3)$$

$$avec \forall(i, j) \in [1, n] \times [1, N], \theta^{(i,j)} = \begin{pmatrix} param_1^{nominal} \\ \vdots \\ param_{i-1}^{nominal} \\ param_i^j \\ param_{i+1}^{nominal} \\ \vdots \\ param_n^{nominal} \end{pmatrix} \quad (IV.4)$$

n correspond aux nombres de paramètres tandis que N correspond au nombre de valeurs discrètes sur l'intervalle de recherche. $\theta^{i,j}$ correspond au vecteur des paramètres où le paramètre i possède la $j^{ème}$ valeur discrète sur l'intervalle de recherche. $U(t)$ est la matrice des sollicitations qui agissent sur le modèle. Soit :

$$\forall(i, j) \in [1, N_{paramètres}] \times [1, N], param_i^j = param_i^{MIN} + (param_i^{MAX} - param_i^{MIN}) \times \frac{j-1}{N-1} \quad (IV.5)$$

Ainsi, pour chaque paramètre, on obtient N courbes de température. Ceci permet de visualiser la variation de la variable de sortie en fonction de la variation d'un paramètre. Les courbes de la figure IV.1 montrent le type de résultats que l'on peut obtenir. L'exemple est donné pour le paramètre C pour quatre jours de jeu de données. Les courbes rouges délimitent la plage de variation parcourue. Elles permettent de calculer l'indice de sensibilité qui correspond à l'écart moyen entre ces deux courbes.

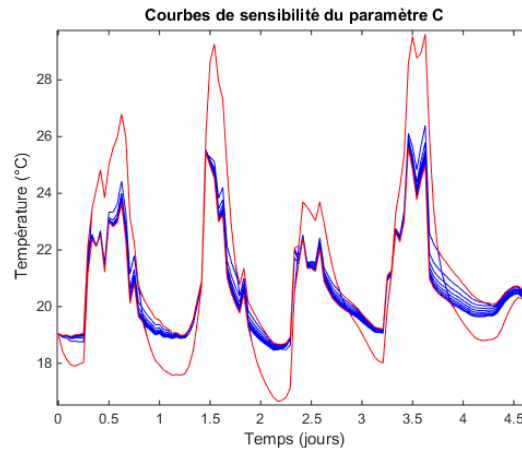


FIGURE IV.1 – Evolution de la variable de sortie pour différentes valeurs du paramètre C

$$\forall i \in [1, n], \text{indice}_{\text{sensibilité}}^i = \frac{1}{n} \sum_{j=1}^N (\max(T_{\text{sensible}}^{(i,j)}) - \min(T_{\text{sensible}}^{(i,j)})) \quad (\text{IV.6})$$

La figure IV.3 rend compte de la sensibilité de chaque paramètre pour des jeux de données expérimentaux répartis sur toute l'année. Un code de couleur (IV.2) permet d'observer où sont situés les paramètres sensibles en moyenne sur un jeu de donnée précis. Les couleurs chaudes montrent les paramètres qui sont les plus sensibles en moyenne tandis que les couleurs froides représentent les paramètres qui ont le moins d'impact sur la variable de sortie.

| Rspace | coefVentil | R1 | C | Residual | R0 | Rcorridor | efficiency | coefOccup | Roffices | RL | fsDir | fsSouth | Tw0 | fsEast | fsWest | coefElec | coefChauff | airLeak |
|--------|------------|----|---|----------|----|-----------|------------|-----------|----------|----|-------|---------|-----|--------|--------|----------|------------|---------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |

FIGURE IV.2 – Légende des paramètres utilisés dans le tableau de sensibilité

| mois | du plus sensible au moins sensible | | | | | | | | | | | | | | | | | | |
|---------|------------------------------------|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|
| janv-15 | 1 | 2 | 6 | 5 | 3 | 4 | 7 | 9 | 11 | 8 | 17 | 12 | 10 | 14 | 18 | 15 | 16 | 13 | 19 |
| févr-14 | 1 | 6 | 5 | 3 | 2 | 4 | 7 | 11 | 12 | 8 | 9 | 14 | 10 | 17 | 13 | 15 | 16 | 18 | 19 |
| févr-15 | 1 | 2 | 3 | 6 | 5 | 9 | 4 | 7 | 11 | 8 | 12 | 14 | 17 | 13 | 15 | 16 | 10 | 18 | 19 |
| avr-14 | 2 | 1 | 4 | 3 | 5 | 6 | 9 | 7 | 8 | 10 | 13 | 12 | 11 | 15 | 16 | 17 | 14 | 19 | 18 |
| juin-14 | 2 | 4 | 3 | 1 | 5 | 8 | 10 | 6 | 13 | 7 | 12 | 9 | 15 | 16 | 11 | 14 | 17 | 19 | 18 |
| juil-14 | 2 | 4 | 1 | 3 | 5 | 6 | 8 | 13 | 9 | 7 | 10 | 12 | 14 | 11 | 15 | 16 | 17 | 19 | 18 |
| août-14 | 2 | 4 | 3 | 1 | 5 | 6 | 8 | 10 | 7 | 13 | 12 | 9 | 14 | 11 | 15 | 16 | 17 | 19 | 18 |
| déc-14 | 1 | 2 | 6 | 5 | 3 | 4 | 7 | 11 | 8 | 9 | 10 | 12 | 14 | 17 | 18 | 15 | 16 | 13 | 19 |
| moyenne | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |

FIGURE IV.3 – Sensibilité locale des paramètres du modèle de PREDIS MHI

On remarque que les paramètres $\text{coef}_{\text{Chauff}}$, $\text{coef}_{\text{Elec}}$ et air_{Leak} sont globalement peu sensibles quel que soit le moment de l'année. Du côté des paramètres sensibles, on remarque que le moment de l'année influe sur la sensibilité. Ainsi, le paramètre R_0 est plutôt sensible en hiver et légèrement moins en été tandis que c'est l'inverse pour le paramètre C .

Afin d'améliorer la convergence de l'algorithme lors de l'identification, on peut fixer certains paramètres. Ainsi, les paramètres les moins sensibles peuvent être fixés à leur valeurs nominales [33]. En effet, réduire le nombre de paramètres peut améliorer la convergence, car comme nous l'avons vu précédemment, le choix de la complexité du modèle influe sur la qualité de l'identification. Les variations des paramètres fixés ayant peu d'impact sur la variable de sortie, la perte en degré de liberté ne devrait pas influencer sur la précision globale des résultats.

Afin d'améliorer les résultats d'identification obtenus précédemment, à l'issue de l'analyse de sensibilité obtenue par la méthode locale, les cinq paramètres les moins sensibles ont été fixés à leur valeur nominale pour les jeux de Juin et de Janvier. A l'issue d'une série d'identifications, les résultats ont été observés et comparés aux résultats obtenus précédemment (figure IV.4).

Nous cherchons à faire gagner le modèle en robustesse, c'est à dire réduire la dispersion de l'erreur à l'issue d'une série d'identifications. Pour cela, nous avons comparé les valeurs d'erreurs obtenues sur 10 identifications pour les deux modèles suivants :

- le modèle initial où tous les paramètres sont estimés ;

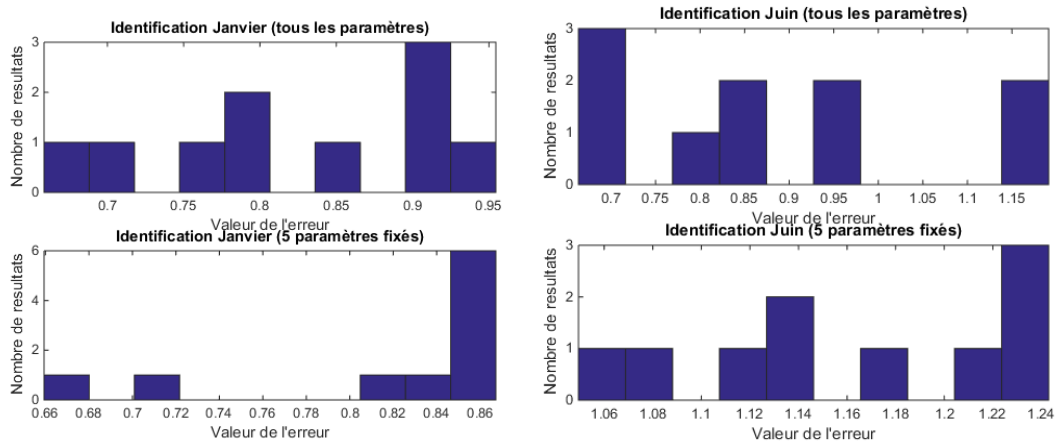


FIGURE IV.4 – Observation de la dispersion de l’erreur pour des modèles avec tous les paramètres et des modèles réduits pour les jeux de données de Janvier et de Juin

- le modèle réduit où les cinq paramètres les moins sensibles sont fixés à leur valeur nominale pendant l’identification, les autres sont estimés.

Les histogrammes des erreurs obtenues sont observables sur les figures IV.4. En abscisse, on observe la valeur d’erreur moyenne tandis que l’ordonnée correspond à l’occurrence de cette valeur d’erreur sur les 10 identifications. Pour le jeu de Janvier, la dispersion du modèle réduit est légèrement réduite mais la meilleure précision reste la même (erreur moyenne environ égale à $0,65^{\circ}C$). Quant au jeu de Juin, la dispersion est approximativement la même, tandis que la précision (à gauche de l’axe des abscisses) est détériorée (on passe d’une erreur moyenne à $0,7^{\circ}C$ à une erreur moyenne à $1,06^{\circ}C$). D’autres modèles réduits avec un nombre de paramètres fixés différents ont été testés et la conclusion reste la même, l’amélioration de la dispersion n’est pas significative pour une précision soit égale, soit détériorée.

Cependant, cette analyse de sensibilité ne prend pas en compte les interactions entre les paramètres et peut donc être limitante pour notre approche. C’est pourquoi nous avons utilisé la méthode de MORRIS.

IV.1.2.ii - b) Méthode de criblage de MORRIS

La méthode de MORRIS a également été testée sur plusieurs jeux de données (deux jeux de données hiver et deux jeux de données été, cf. figure IV.5).

Parmi les paramètres sensibles, on retrouve $coef_{Ventil}$, R_1 et R_{space} qui faisaient partie des paramètres sensibles avec la méthode locale. Ici, T_{w0} et $efficiency$ sont également sensibles. Les paramètres non sensibles sont les coefficients de flux solaire, $coef_{Occup}$, $coef_{Elec}$, R_0 et C . $coef_{Chauff}$ est sensible en hiver et non sensible en été.

Comme précédemment, les paramètres les moins sensibles à l’issue de l’analyse de sensibilité ont été fixés à leurs valeurs nominales. Les modèles réduits ont été identifiés sur des jeux de Janvier et de Juin et les résultats ont été comparés au modèle initial. Sur les figures IV.6 et IV.7, on peut observer l’impact de la réduction du modèle sur les résultats d’erreur. Ainsi, on remarque que pour le jeu d’hiver, la dispersion de l’erreur est toujours grande mais

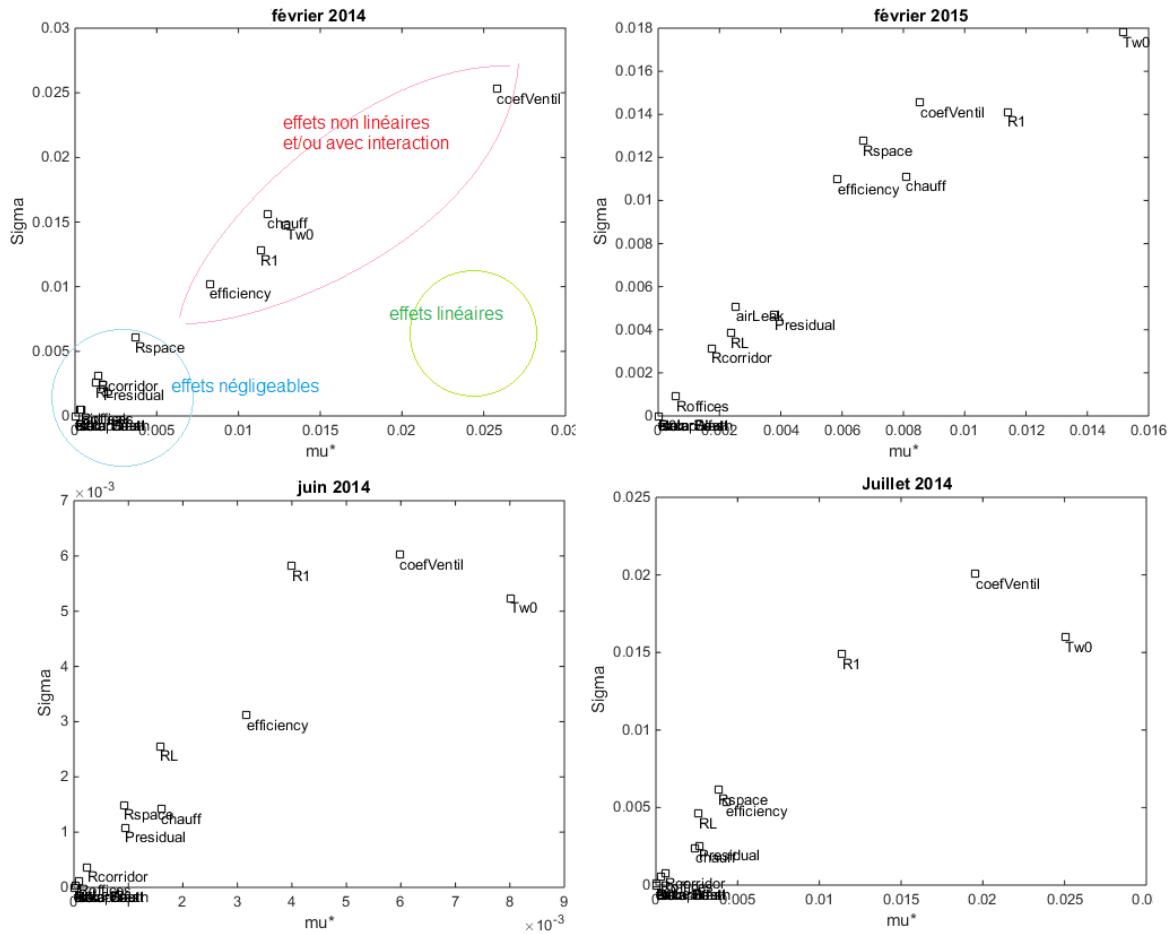


FIGURE IV.5 – Analyse de sensibilité des paramètres du modèle pour quatre jeux de données avec la méthode de MORRIS

légèrement meilleure en précision que celle du modèle initial. La courbe de la température intérieure obtenue à l'issue de l'identification paraît également plus lisse.

Par ailleurs, concernant le jeu de Juin, l'identification est beaucoup moins précise. En effet, la valeur de l'erreur pour la meilleure identification du modèle réduit (environ égale à 1.25°C) est moins bonne que la valeur d'erreur pour la pire identification du modèle initial (environ égale à 1°C).

Ainsi, de même que précédemment, la réduction du modèle selon l'analyse de sensibilité de MORRIS ne montre pas d'amélioration significative par rapport à l'identification du modèle.

La réduction du nombre de paramètres à identifier à l'aide de l'analyse de sensibilité n'améliorant pas la dispersion globale (paramétrique ou du vecteur d'erreur), il peut être intéressant de mieux étudier ces aspects. En effet, ils traduisent un problème d'identifiabilité. L'arrivée en butée de certaines valeurs de paramètres identifiés est également un critère intéressant à étudier. Si l'espace de recherche des paramètres est défini avec des bornes larges, les valeurs arrivées en butée n'auront plus aucun sens physique.

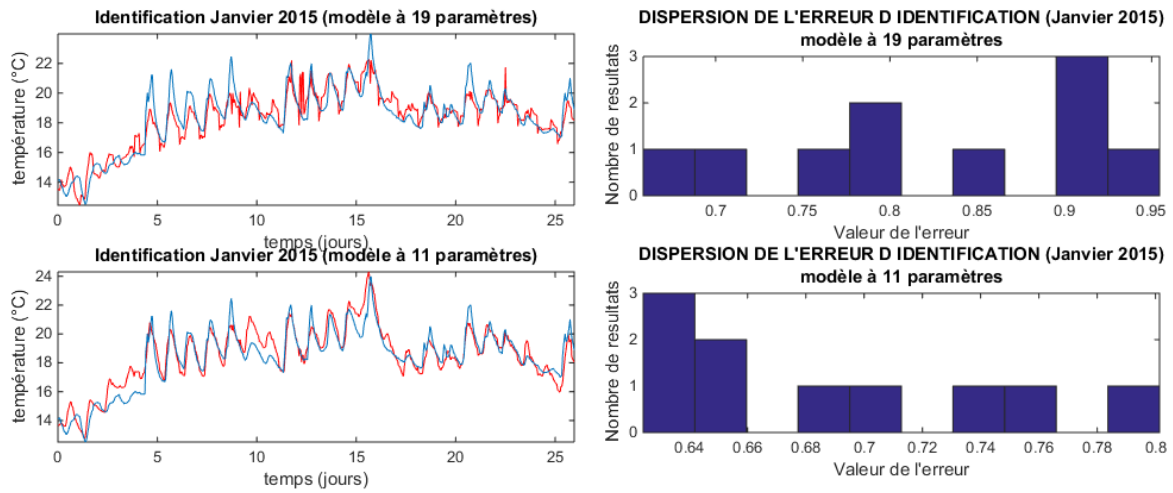


FIGURE IV.6 – Meilleur résultat d'identification et observation du comportement de l'erreur à l'issue d'une série d'identifications pour le modèle initial (à 19 paramètres) et un modèle réduit (à 11 paramètres) pour un jeu de Janvier 2015

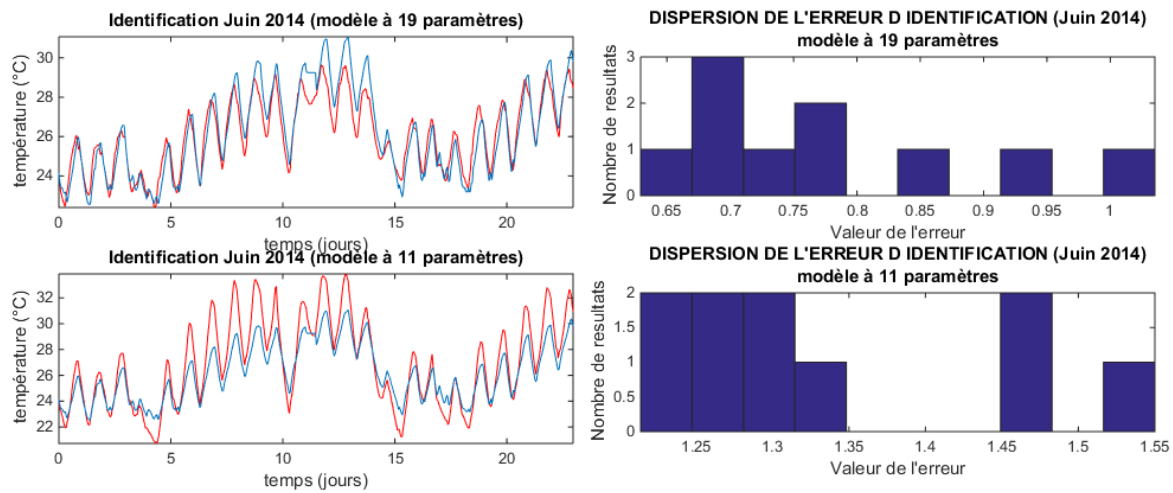


FIGURE IV.7 – Meilleur résultat d'identification et observation du comportement de l'erreur à l'issue d'une série d'identifications pour le modèle initial (à 19 paramètres) et un modèle réduit (à 11 paramètres) pour un jeu de Juin 2014

IV.1.3 Analyse de la dispersion paramétrique et de l'arrivée en butée des valeurs identifiées

En réalisant une série d'identifications, on peut observer le comportement des valeurs des paramètres.

La figure IV.8 montre les valeurs de paramètres obtenues pour 10 identifications réalisées sur un jeu de données de Juin 2014 avec des bornes de recherche larges. Certains paramètres sont très dispersés ($coef_{elec}$, $coef_{chauff}$, C , air_{Leak}) et d'autres arrivent constamment en butée de leur domaine de validité ($R_{offices}$, $R_{corridor}$, les coefficients de flux solaires).

Un phénomène de dispersion sur l'espace de recherche des paramètres traduit un problème d'identifiabilité. Si ce phénomène est récurrent quel que soit le jeu de données utilisé alors

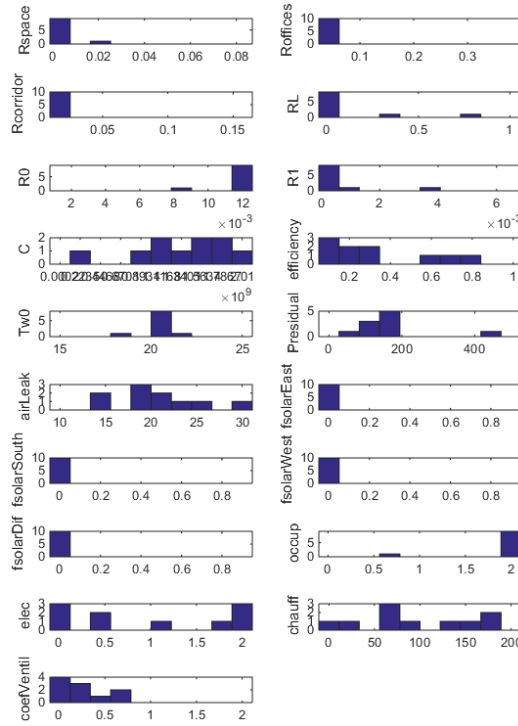


FIGURE IV.8 – Observation de la dispersion paramétrique pour des bornes larges pour un jeu de données de Juin 2014

cela signifie que peu importe les sollicitations, le paramètre est difficilement identifiable. Le manque d'identifiabilité peut être dû à plusieurs facteurs et en particulier :

- Un mauvais choix dans la complexité du modèle ;
- Des signaux trop pauvres qui contiennent peu d'informations.

A l'aide de plusieurs analyses de sensibilité, nous avons essayé de réduire la complexité du modèle, sans succès. Nous avons donc réalisé des expérimentations afin d'obtenir des signaux plus riches.

IV.2 Étude de la richesse spectrale des signaux d'entrée

L'excitation des températures des salles PREDIS est liée à la température extérieure qui a un comportement basse fréquence avec entre autres une raie spectrale 1/24h. Ce genre d'excitation est très limitée, d'autant plus qu'elle est atténuée en propagation par l'isolation de la structure.

Les apports internes influent également sur la température intérieure mais une fois de plus, ils ne sollicitent généralement pas suffisamment les pièces pour observer des changements brusques dans la température. Sur la figure IV.9, on peut observer le spectre de la température extérieure et des températures des salles PREDIS MHI pour un jeu de Janvier 2015. Sur ces signaux, la valeur moyenne de la température a été retirée afin d'observer le comportement périodique des températures. On peut remarquer une raie spectrale à 1/24h qui traduit la

tendance journalière des pièces et qui est dû au comportement de la température extérieure. Les spectres des températures intérieures comportent également une raie spectrale à $1/504h$ qui correspond à une période de trois semaines. Les pièces sont occupées, chauffées et ventilées en semaine tandis que ce n'est pas le cas le week end, nous devrions donc observer une périodicité hebdomadaire avec une raie spectrale à $1/168h$. En réalité, celle-ci est quasiment confondue avec la raie à $1/504h$ car la durée du jeu de données ne doit pas être suffisante pour observer la périodicité.

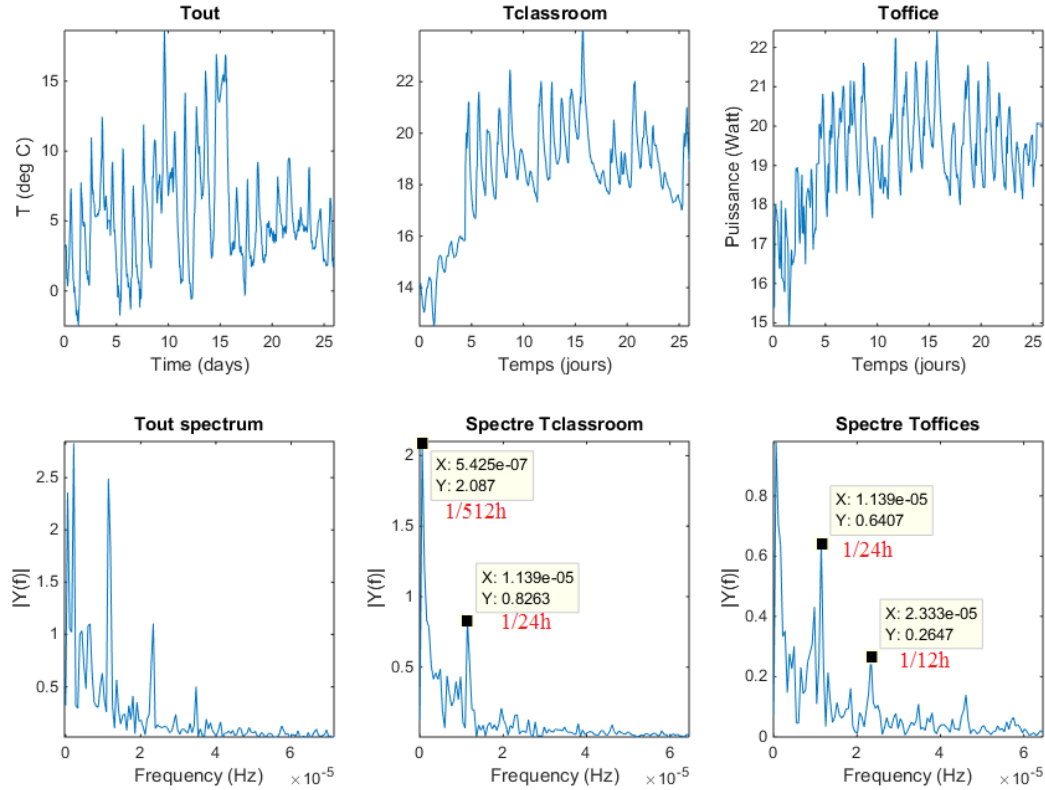


FIGURE IV.9 – Signaux temporels et fréquentiels de la température extérieure et des deux salles de PREDIS MHI

Afin de posséder les informations nécessaires à une bonne identification paramétrique, un signal doit être excité par ce qui se rapprochera le plus d'un bruit blanc, qui excite toutes les fréquences. On utilise généralement une excitation par séquence binaire pseudo aléatoire (SBPA) pour exciter les signaux d'entrée du modèle et en tirer des informations pertinentes pour l'identification.

Saint Gobain [30] a également récemment montré qu'un bâtiment pouvait être entièrement caractérisé par un échelon de température. Cette expérimentation nécessite la mise à disposition du bâtiment vide pendant 3 jours. Il s'agit ici de caractériser le bâtiment dans son ensemble, de manière globale et sans distinction de saison. Dans notre cas, nous cherchons plutôt simple et facilement recalibrable qui aurait une bonne capacité d'extrapolation.

Afin d'obtenir des signaux riches, nous avons piloté la pompe à chaleur de l'espace bureau à l'aide d'une séquence binaire pseudo aléatoire afin d'enregistrer des signaux d'entrée contenant plus d'information spectrale. Cette séquence durait 15 jours.

IV.2.1 Sollicitation de la pompe à chaleur de l'espace bureau à l'aide d'une séquence binaire pseudo-aléatoire

IV.2.1-i Séquence binaire pseudo aléatoire

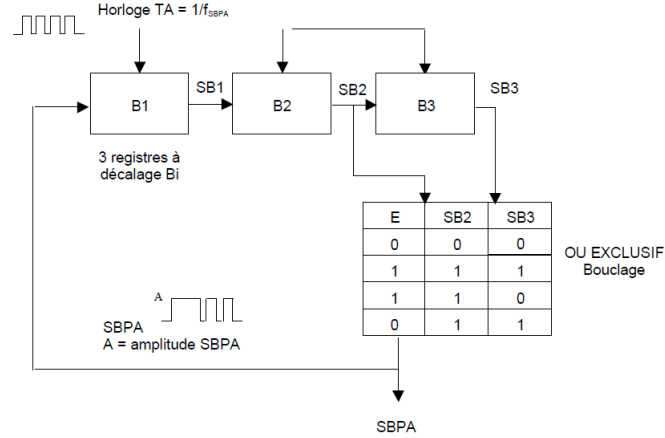


FIGURE IV.10 – Génération d'une SBPA par registre à décalage bouclé

La SPBA est un signal pseudo aléatoire car elle est caractérisée par une longueur de séquence à l'intérieur de laquelle la largeur des impulsions varie aléatoirement, mais, sur un grand horizon de temps, ces variations sont périodiques, la période étant définie par la longueur de la séquence. La SBPA est générée à l'aide de registres à décalage (réalisés en matériel ou logiciel) bouclés. La longueur d'onde maximale d'une séquence est $2^N - 1$ où N est le nombre de cellules du registre à décalage. La figure IV.10 montre l'exemple d'une SBPA de longueur $7 = 2^3 - 1$ obtenue à l'aide d'un registre à décalage ayant 3 cellules.

IV.2.1-ii Expérimentation sur les salles PREDIS MHI : pilotage de la pompe à chaleur à l'aide d'une SBPA

Afin d'enrichir les jeux de données d'entrée servant à l'identification des salles PREDIS MHI, la pompe à chaleur de l'espace bureau a été pilotée par une séquence binaire pseudo aléatoire. Ce genre de pilotage est disponible grâce au système de monitoring et de contrôle que nous avons installé dans la plateforme PREDIS MHI en collaboration avec la société Vesta-System. Cette expérimentation a été réalisée pendant l'été 2014 sur une période d'un mois et demi (soit trois séquences à la suite).

Une analyse spectrale permettra de vérifier l'enrichissement des données. Une série d'identifications réalisées sur deux jeux de données différents, l'un contenant des signaux excités par la pompe à chaleur et l'autre non, pourra permettre de juger de l'amélioration du procédé d'identification.

IV.2.1.ii - a) Analyse spectrale des jeux de données

La courbe IV.11 rend compte du fonctionnement de la pompe à chaleur pilotée par la séquence binaire pseudo aléatoire.

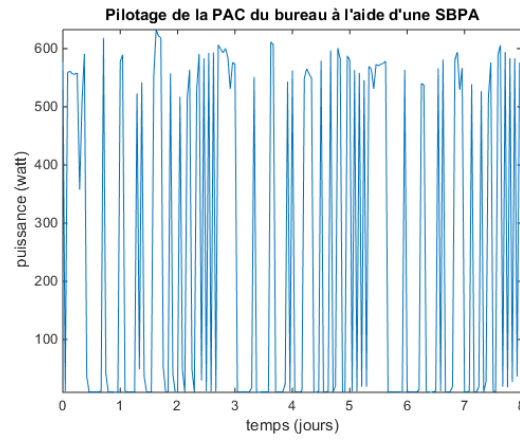


FIGURE IV.11 – Pilotage de la pompe à chaleur de l'espace bureau à l'aide d'une séquence binaire pseudo aléatoire

Cette expérimentation a débuté le 15 Juillet et s'est terminée fin Août. Afin de juger de l'enrichissement des jeux de données, il faut comparer deux jeux de données subissant approximativement les mêmes sollicitations. Comme le laboratoire est fermé les deux premières semaines d'août, les deux jeux de données retenus pour cette comparaison ont été collectés entre le 7 et le 15 Juillet et entre le 18 et le 26 Août. A ces dates, la salle de classe est vide tandis que l'espace bureau est occupé.

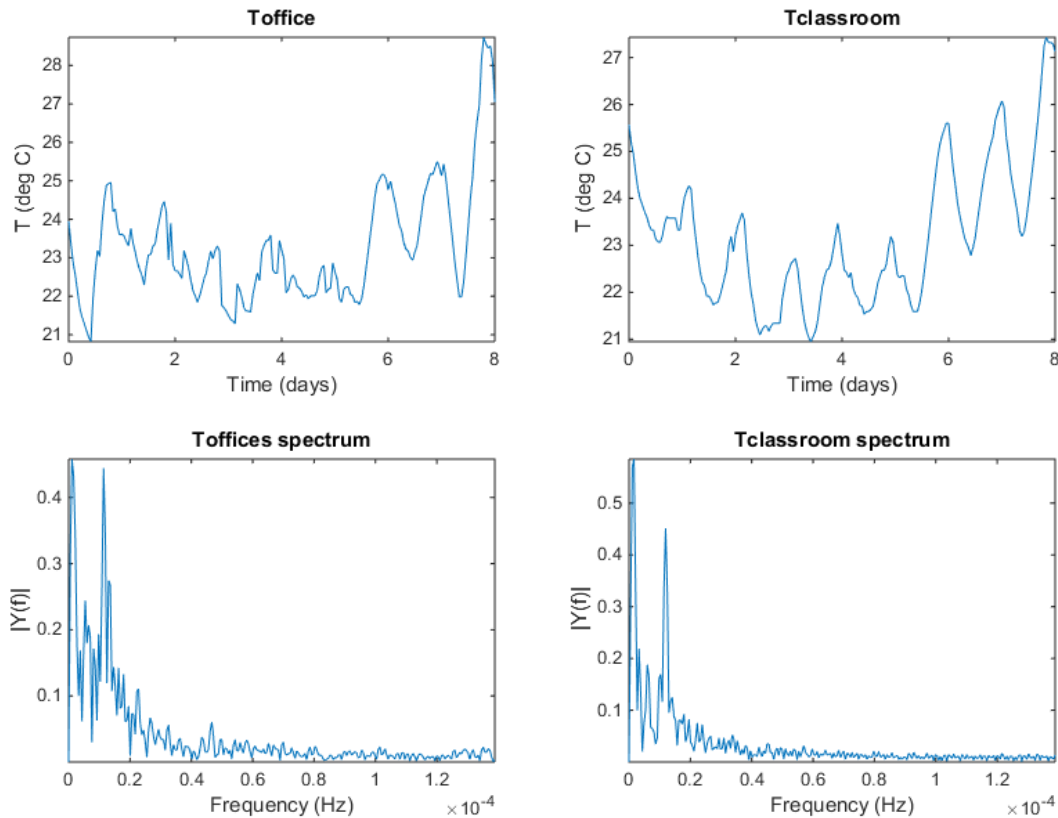


FIGURE IV.12 – Spectre de la température de la salle de classe et de l'espace bureau pour un jeu de Juillet 2014, non excité par une séquence binaire pseudo aléatoire

Sur le spectre du jeu de données collecté en Juillet (figure IV.12), on remarque les mêmes raies spectrales que précédemment à $1/24h$ et $1/168h$ qui traduisent la périodicité quotidienne

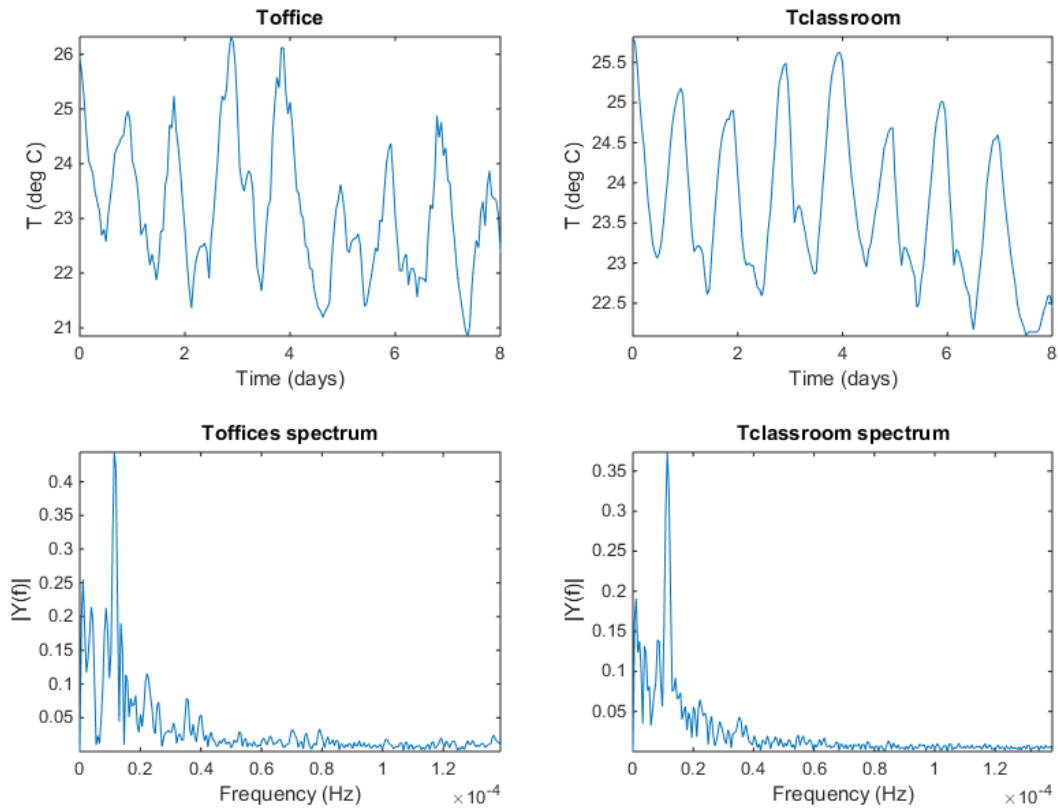


FIGURE IV.13 – Spectre de la température de la salle de classe et de l'espace bureau pour un jeu d'août 2014, excité par une séquence binaire pseudo aléatoire

et hebdomadaire liée à l'occupation (cette dernière correspond à la différence de comportement semaine/ week end).

Si on observe le spectre du jeu de données collecté au mois d'août (figure IV.13) pendant l'expérimentation pour l'enrichissement des jeux de données, on remarque que la raie 1/168h a disparu. En effet, le scénario de la pompe à chaleur masque les autres sollicitations internes, et la différence semaine / week-end devient beaucoup moins flagrante, puisque le scénario ne s'arrête pas le week end.

IV.2.1.ii - b) Identification des jeux de données

La température de la salle de classe est moins impactée par le pilotage de la pompe à chaleur. Cependant, la température de l'espace bureau qui a été excitée par le pilotage de la pompe à chaleur devrait avoir un contenu spectral plus dense et constitue donc une entrée pertinente pour l'identification de la salle de classe.

Les identifications suivantes ont été réalisées :

- Une série de 10 identifications de la salle de classe pour le mois de Juillet 2014 ;
- Une série de 10 identifications de l'espace bureau pour le mois de Juillet 2014 ;
- Une série de 10 identifications de la salle de classe pour le mois d'août 2014 ;
- Une série de 10 identifications de l'espace bureau pour le mois d'août 2014.

A travers ces identifications nous observerons et comparerons :

- La dispersion de l'erreur ;
- Le comportement de l'erreur (répartition, moyenne, maximum atteint) ;
- La dispersion paramétrique ;
- Le nombre de paramètres en limite de validité ;
- La qualité de la validation sur un autre jeu de donnée d'été (Juin 2014).

Tous ces critères nous permettront de juger de l'amélioration de l'identification ou non.

A l'issue de la série d'identifications, le meilleur modèle est observé. Le tableau IV.1 rend compte des erreurs moyennes et maximales obtenues pour les deux salles en identification pour les deux jeux de données et en validation sur le jeu de données de Juin 2014.

| | | Août | | Juillet | |
|------------------------|----------------|--------|-----------------|---------|-----------------|
| | | bureau | salle de classe | bureau | salle de classe |
| Identification | erreur moyenne | 0,68 | 0,43 | 0,80 | 0,60 |
| | erreur maximum | 2,24 | 1,25 | 2,70 | 2,03 |
| Validation (Juin 2014) | erreur moyenne | 0,95 | 0,84 | 1,00 | 0,92 |
| | erreur maximum | 2,66 | 3,12 | 2,99 | 3,12 |

TABLEAU IV.1 – Comparaison des valeurs d'erreurs en identification et en validation pour les deux jeux de données collectés pendant l'été 2014, l'espace bureau excité par une SBPA et la salle de classe non

Les résultats obtenus pour le jeu d'Août qui a un contenu spectral plus riche sont meilleurs mais restent proche des résultats de Juillet.

Concernant la dispersion de l'erreur, l'amélioration n'est pas flagrante. Si on regarde les figures IV.14, on peut remarquer que concernant le bureau, il n'y a pas d'amélioration de la précision et l'erreur est même plus dispersée avec un écart type de 0,0857 pour la série d'identifications d'Août contre 0,0548 pour la série d'identifications de Juillet (cf. tableau IV.2). Pour la salle de classe, c'est l'inverse, l'écart type est bien deux fois plus petit pour le jeu d'Août que pour le jeu de Juillet. De plus, la précision est améliorée.

| | Juillet 2014 | Août 2014 |
|-----------------|--------------|-----------|
| salle de classe | 0,10 | 0,05 |
| espace bureau | 0,05 | 0,09 |

TABLEAU IV.2 – Valeur des écarts types des vecteurs d'erreurs

La dispersion paramétrique est quantifiée par la moyenne des écarts types des valeurs de paramètres identifiés divisés par la plage de recherche des paramètres afin d'obtenir des résultats normalisés.

Sur le tableau IV.3, on peut voir qu'elle est à peu près identique pour les deux jeux de données concernant la salle de classe tandis qu'elle est légèrement meilleure pour le mois d'Août par rapport à l'espace bureau.

Cette expérimentation n'a pas donné des résultats très concluants. Les résultats d'identification de la salle de classe sont légèrement meilleurs pour le jeu d'Août (plus riche) que pour le jeu de Juillet mais il est difficile de lier cela avec la richesse spectrale des jeux de données.

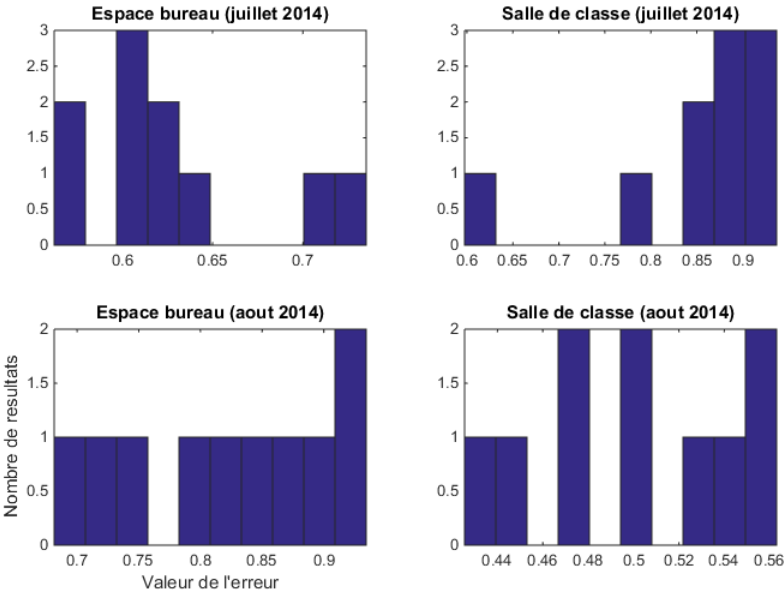


FIGURE IV.14 – Comparaison des vecteurs d’erreurs obtenues à l’issue d’une série d’identification pour l’espace bureau et la salle de classe en Juillet et Août 2014

| | Juillet 2014 | Août 2014 |
|-----------------|--------------|-----------|
| salle de classe | 0,18 | 0,18 |
| espace bureau | 0,34 | 0,21 |

TABLEAU IV.3 – Valeur moyenne des écarts types des vecteurs de valeurs de paramètres identifiées

IV.2.2 Longueur des données nécessaires à l’identification

La longueur des signaux utilisés joue également sur la qualité de l’identification et de la prédiction. A l’aide du jeu de données simulé sur une année issu du logiciel COMIFIE-Pléiade, des tests d’identification en fonction de la longueur du jeu d’identification et de validation ont été réalisés.

Le tableau IV.4 montre les résultats d’erreur moyenne et de dispersion obtenus en fonction de la longueur du jeu de données servant à l’identification. Si on regarde ces chiffres sur une courbe (cf. IV.15), on remarque que les résultats ont tendance à s’améliorer. Néanmoins, au delà de 30 jours de jeu de données pour l’identification, les résultats se détériorent.

| | | erreur moyenne | dispersion paramétrique moyenne | dispersion de l’erreur |
|------------------------|-----------|-------------------|------------------------------------|---------------------------|
| Durée d’identification | 1 jours | 0,60 | 0,22 | 0,76 |
| | 5 jours | 0,40 | 0,21 | 0,62 |
| | 10 jours | 0,45 | 0,12 | 0,61 |
| | 30 jours | 0,20 | 0,12 | 0,31 |
| | 100 jours | 0,48 | 0,14 | 0,57 |

TABLEAU IV.4 – Observation de l’erreur moyenne et des dispersions obtenues en fonction de la longueur du jeu d’identification

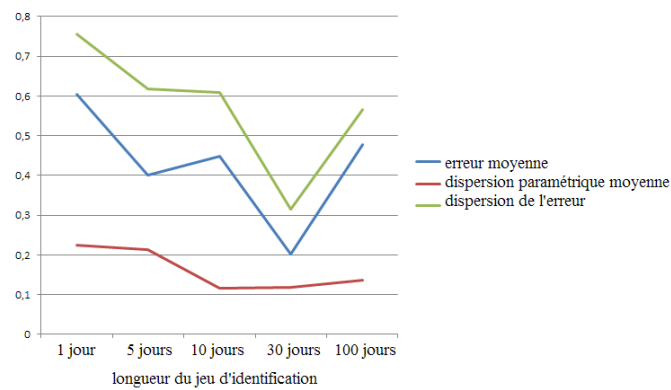


FIGURE IV.15 – Observation de l'erreur moyenne et des dispersions obtenues en fonction de la longueur du jeu d'identification

Des tests de validation ont également été réalisés. Dans le tableau IV.5 nous pouvons observer les erreurs moyennes obtenues en fonction de la longueur du jeu de données ayant servi à l'identification et de la longueur du jeu de données pour la validation.

| | | Validation (erreur moyenne) | | | | |
|------------------------|-----------|-----------------------------|---------|----------|----------|-----------|
| | | 1 jour | 5 jours | 10 jours | 30 jours | 100 jours |
| Durée d'identification | 1 jour | 0,39 | 0,69 | 0,75 | 0,77 | 0,77 |
| | 5 jours | 0,67 | 0,57 | 0,54 | 0,48 | 0,95 |
| | 10 jours | 0,29 | 0,23 | 0,24 | 0,23 | 0,28 |
| | 30 jours | 0,18 | 0,25 | 0,23 | 0,22 | 0,38 |
| | 100 jours | 0,25 | 0,34 | 0,33 | 0,31 | 0,55 |

TABLEAU IV.5 – Observation de l'erreur moyenne en validation en fonction de la longueur du jeu d'identification et du jeu de validation

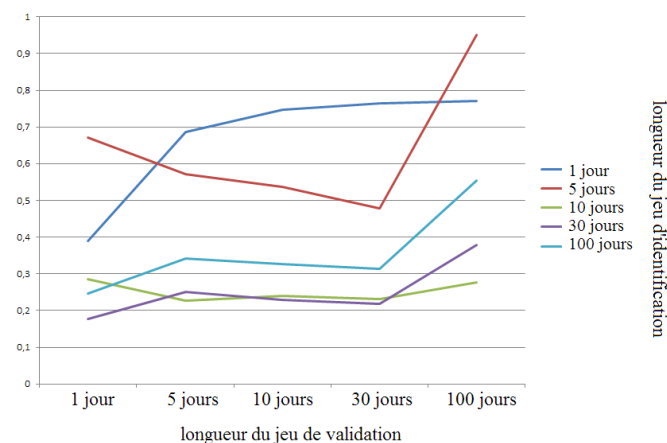


FIGURE IV.16 – Observation de l'erreur moyenne en fonction de la longueur du jeu d'identification et du jeu de validation

Ainsi, sur la figure IV.16, on peut voir que si l'on souhaite prédire la température intérieure à un jour, un jeu de donnée ayant une longueur comprise entre 30 et 100 jours donnera les meilleurs résultats. Globalement, quelle que soit la longueur du jeu à prédire, les meilleurs résultats sont obtenus avec des jeux de données d'identification compris entre 10 et 30 jours (courbes verte et violette).

On peut donc remarquer qu'un signal très court contient très peu d'information, ou en tout cas pas suffisamment pour obtenir de bons résultats en validation. A l'inverse, un signal trop long contient des événements trop différents pour bien caractériser un jeu de donnée en validation.

Ceci traduit deux choses importantes :

- La durée du jeu servant à l'identification est un élément à ne pas négliger dans le cadre de la gestion anticipative et dépend de :
 - l'horizon recherché ;
 - la richesse spectrale des jeux de données ;
- La recalibration du modèle au fur et à mesure de l'avancement de l'année est primordiale.

Concernant le premier point, la longueur des jeux de données à utiliser pour obtenir les meilleurs résultats dépend du modèle et du contenu spectral des signaux. Ici, l'expérience a été réalisée avec des signaux qui ont été obtenus par simulation à l'aide du logiciel COMFIE-Pléiade. Ce sont des signaux relativement pauvres spectralement mais très bien connus (scénarios d'occupation, de chauffage, de ventilation). Avec des signaux expérimentaux, l'étude est plus complexe car nous devons faire avec les incertitudes, les bruits de mesures et les erreurs de capteurs (valeurs manquantes et aberrantes).

Quant au deuxième point, cette expérience nous montre en effet qu'il est pertinent de ré-identifier notre modèle assez régulièrement puisqu'une identification avec un jeu d'une centaine de jours ne donne pas des résultats concluants. En effet, on pourrait penser que plus le jeu est grand, mieux il sera à même de représenter toutes sortes de situations. Cependant, bien que donnant une bonne idée du comportement global du système, il représentera moins bien les comportements ponctuels. Ceci peut s'expliquer par le fait que notre modèle contient des simplifications (apports solaires, ouvrants, ...) et des approximations (notamment sur les phénomènes dynamiques et la discrétisation de la ventilation). Ainsi, l'ajustement des paramètres en fonction des signaux expérimentaux permet d'absorber les éventuelles incertitudes et les différences liées aux comportements été et hiver.

C'est pourquoi il est important de choisir une longueur d'identification idéale pour la gestion anticipative. Il faut que cette période de temps possède les qualités suivantes :

- Contenir assez d'informations quant à la saison en cours ;
- Permettre un recalibrage rapide et robuste ;

Une longueur suffisante permet également de s'affranchir des petits phénomènes qui peuvent survenir comme l'ouverture des portes, un changement brutal de ventilation, une occupation de quelques personnes inattendues, etc. En effet, sur un jeu de données trop petit, ces phénomènes peuvent avoir plus d'impact et influencer l'identification. Ainsi, la difficulté de l'identification peut également être liée aux phénomènes mal maîtrisés qui existent dans les jeux de données mais ne sont pas nécessairement modélisés (comme l'état d'ouverture des

portes notamment). Ces phénomènes ont été simulés à l'aide de logiciels thermiques et les jeux de données obtenus ont servi pour réaliser des tests en identification avec notre modèle.

IV.2.3 Pertinence du choix des bornes sur les résultats d'estimation paramétrique

La largeur des intervalles de recherches a une influence non négligeable sur les résultats du modèle de gestion. En effet, celui ci a été choisi physique et c'est pourquoi il est logique de définir les bornes du modèle en fonction des incertitudes que nous avons sur les valeurs analytiques des paramètres qui ont été calculées en fonction des caractéristiques du bâtiment.

Plus les intervalles sont grands et plus le modèle aura de liberté et plus il convergera vers une erreur faible en identification. C'est ce que nous avons pu observer en définissant un coefficient d qui divise ou multiplie les valeurs nominales des paramètres pour obtenir les bornes de l'espace de recherche. Nous avons fait varier la valeur de ce coefficient d afin d'observer l'évolution de l'erreur obtenue en identification. Ainsi plus le coefficient d est grand et plus l'erreur est faible, comme nous pouvons l'observer sur la figure IV.17.

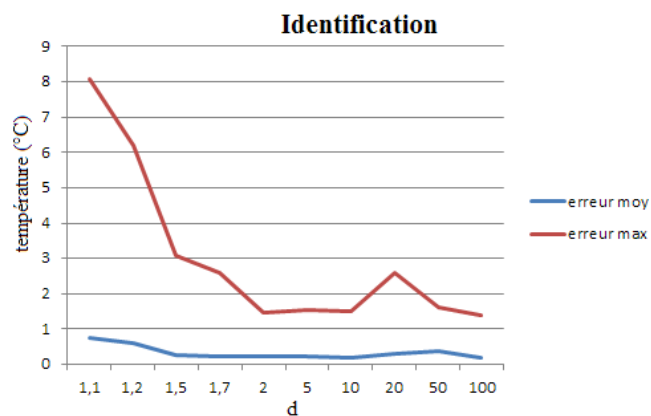


FIGURE IV.17 – Évolution de l'erreur d'identification en fonction de la largeur des bornes

Cependant, plus la valeur du coefficient d est grande et moins la cohérence physique du modèle est conservée. Les figures IV.18 montrent les résultats obtenus en validation sur plusieurs autres mois après que le modèle ait été identifié en janvier (données COMFIE Pléiades) avec plusieurs bornes pour l'espace de recherche.

Les résultats en terme d'erreurs moyennes sont à peu près équivalents, mais on remarque quand même une légère détérioration de l'erreur lorsque le coefficient d augmente.

Il s'agit de trouver un compromis intéressant en fonction des incertitudes que nous avons sur les calculs des paramètres tout en laissant une certaine marge de manoeuvre à l'algorithme pour converger vers une erreur satisfaisante en identification et en validation.

Tandis que le tableau III.5 du chapitre III présentait les intervalles liés aux incertitudes sur les calculs analytiques, le tableau IV.6 décrit les intervalles de recherches qui ont finalement été retenus pour le modèle.

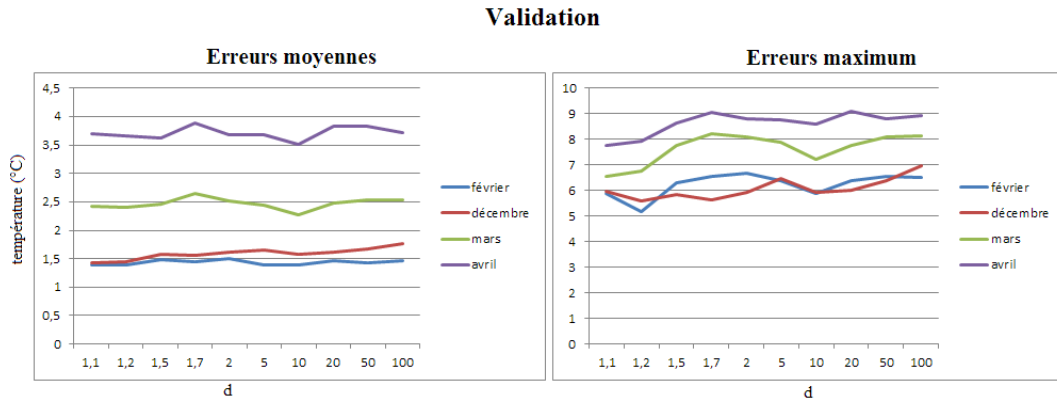


FIGURE IV.18 – Évolution des erreurs de validations en fonction de la largeur des intervalles

| | valeur minimum | valeur nominale | valeur maximum |
|---------------------|-------------------|--------------------|--------------------|
| R_{space} | $R_{space}/5$ | 0,017 | $R_{space} * 5$ |
| $R_{offices}$ | $R_{offices}/3$ | 0,13 | $R_{offices} * 3$ |
| $R_{corridor}$ | $R_{corridor}/3$ | 0,052 | $R_{corridor} * 3$ |
| R_L | $R_L/10$ | 0,1007 | $R_L * 10$ |
| R_0 | $R_0/3$ | 0,004 | $R_0 * 3$ |
| R_1 | $R_1/5$ | 0,0013 | $R_1 * 5$ |
| C | $2,01e^5$ | $2,01e^7$ | $2,01e^9$ |
| $efficiency$ | 0,1 | 0,8 | 0,99 |
| T_{w0} | 15 | 20 | 25 |
| $P_{residual}$ | 0 | 200 | 1000 |
| air_{Leak} | 10 | 22 | 30 |
| $coef_{occup}$ | 0,1 | 2 | 4 |
| $coef_{elec}$ | 0,1 | 1 | 3 |
| $coef_{chauff}$ | 0,1 | 1 | 200 |
| $coef_{ventil}$ | 0,1 | 1 | 2 |
| f_{solar}^{East} | 0 | 0,3 | 0,9 |
| f_{solar}^{South} | 0 | 0,3 | 0,9 |
| f_{solar}^{West} | 0 | 0,3 | 0,9 |
| f_{solar}^{Diff} | 0 | 0,3 | 0,9 |

TABLEAU IV.6 – Valeur nominales et intervalles de recherches pour le modèle PREDIS MHI

Ces bornes souples et adaptées à chaque paramètres nous donnent de très bons résultats d'identification et de validation en comparaison d'un modèle dont les bornes seraient choisis globalement et de façon arbitraire.

IV.3 Incertitudes sur les phénomènes

De nombreux phénomènes peuvent survenir sans avoir été anticipés et peuvent perturber l'identification et la prédiction. L'état d'ouverture des portes et l'influence des scénarios de ventilation notamment, ont été simulés à partir du logiciel IDA-ICE et ces jeux de données ont été identifiés afin de mieux observer l'impact de phénomènes non maîtrisés sur l'identification.

IV.3.1 Influence de l'état d'ouverture des portes sur l'identification de la température intérieure de PREDIS MHI

La gestion anticipative nécessite des modèles simples et facilement recalibrables. Ainsi, certains aspects du système à modéliser ont dû être négligés comme l'état d'ouverture des portes par exemple. A l'aide du logiciel IDA-ICE dans lequel ont été modélisées les salles PREDIS MHI, deux cas de scénarios d'ouverture de portes ont été testés :

- Les portes sont tout le temps ouvertes ;
- Les portes s'ouvrent et se ferment selon un scénario aléatoire (censé reproduire un comportement normal).

L'identification de ces cas avec notre modèle permettra de quantifier l'influence de l'état d'ouverture des portes sur la température intérieure et si cet aspect est réellement négligeable dans la modélisation.

Dans la modélisation simplifiée des salles PREDIS MHI, seule la résistance $R_{corridor}$ est liée au renouvellement d'air dû à l'ouverture des portes dans le couloir. Dans le cas où les portes sont tout le temps ouvertes, il est logique de rechercher une résistance constante puisque la quantité de fuite d'air dans les salles PREDIS MHI ne devrait pas changer au cours du temps. Il s'agit, en comparant les deux scénarios, de vérifier que l'ouverture et la fermeture des portes ne vient pas perturber l'identification du paramètre $R_{corridor}$ et du modèle en général.

| | Scénario : portes ouvertes | | | |
|----------------|--------------------------------|-------------------|---------------------------------|---------------------------------------|
| | erreur moyenne | dispersion erreur | dispersion paramétrique moyenne | dispersion coefficient $R_{corridor}$ |
| Janvier | 0,20 | 0,22 | 0,22 | 0,04 |
| Février | 0,21 | 0,37 | 0,20 | 0,01 |
| Juin | 0,41 | 0,27 | 0,29 | $1.8e^{-18}$ |
| Juillet | 0,58 | 0,16 | 0,18 | $1.8e^{-18}$ |
| moyenne | 0,35 | 0,25 | 0,22 | 0,01 |
| | Scénario : ouverture aléatoire | | | |
| | erreur moyenne | dispersion erreur | dispersion paramétrique moyenne | dispersion coefficient air_{Leak} |
| Janvier | 0,24 | 0,16 | 0,22 | 0,09 |
| Février | 0,27 | 0,41 | 0,24 | 0,23 |
| Juin | 0,68 | 0,18 | 0,17 | 0,002 |
| Juillet | 0,73 | 0,16 | 0,20 | $1.36e^{-4}$ |
| moyenne | 0,48 | 0,23 | 0,21 | 0,08 |

TABLEAU IV.7 – Comparaison des résultats d'identification pour des scénarios d'ouverture de porte à divers moments de l'année

Pour plusieurs jeux de données à divers moments de l'année, une série d'identification a été réalisée pour les deux scénarios d'ouverture de portes. Les résultats obtenus sont résumés dans le tableau IV.7. Ainsi on remarque que le fait que les portes soient ouvertes de manière aléatoire n'a pas d'impact sur la dispersion globale mais la dispersion de la résistance thermique du couloir est plus grande quand les portes sont ouvertes de manière aléatoire. De

plus, l'erreur moyenne est légèrement dégradée pour le scénario où les portes sont ouvertes de manière aléatoire mais celle-ci reste acceptable dans le cadre de la gestion anticipative.

Ainsi, l'influence de l'ouverture des portes a un impact négligeable sur le modèle.

IV.3.2 Influence des scénarios de ventilation sur l'identification des salles PREDIS MHI

Au début du chapitre III, nous avons vu que lors d'un changement brutal de scénario sur la ventilation dans le jeu de validation et malgré le fait que ce scénario ne soit pas représenté dans le jeu servant à l'identification, les résultats étaient malgré tout relativement bons. Comme évoqué précédemment, le modèle ayant une cohérence physique et la ventilation, discrétisée en quatre paliers, étant représentée dans le modèle, il est logique que la validation donne de bons résultats. Néanmoins, la valeur d'erreur à ce moment là est moins bonne que celle qui correspond au scénario qui était présent dans l'identification. Ainsi, il est intéressant de voir à quel point les scénarios de ventilation peuvent influencer l'identification et la validation. Pour cela, plusieurs scénarios de ventilation ont été simulés dans le logiciel COMFIE-Pléiade et les jeux de données issus de ces simulations ont servi à l'identification.

Nous avons identifié et validé des jeux de données ayant les caractéristiques suivantes :

- Scénario de ventilation standard : $420\text{m}^3/\text{h}$ de débit d'air entre 8h et 18h tous les jours de semaine ;
- Pas de ventilation (période de vacances scolaire) ;
- Scénario mixte contenant des jours de ventilation standard et des jours de vacances.

| | Erreur moyenne | Dispersion de l'erreur | Dispersion paramétrique |
|---------------------------|----------------|------------------------|-------------------------|
| Scénario avec ventilation | 0,24 | 0,04 | 0,16 |
| Scénario sans ventilation | 0,20 | 0,14 | 0,21 |
| Scénario mixte | 0,23 | 0,63 | 0,22 |

TABLEAU IV.8 – Observation des résultats d'identification pour divers scénario de ventilation

Dans le tableau IV.8, on remarque que les résultats d'identification sont proches malgré les différences de scénarios mais le vecteur d'erreur obtenu dans le cas du scénario mixte est relativement dispersé comparé aux deux autres scénarios.

| | | Validation (°C) | | |
|----------------|---------------------------|---------------------------|---------------------------|----------------|
| | | Scénario avec ventilation | Scénario sans ventilation | Scénario mixte |
| Identification | Scénario avec ventilation | 0,26 | 0,82 | 0,47 |
| | Scénario sans ventilation | 1,24 | 0,34 | 0,66 |
| | Scénario mixte | 0,41 | 0,48 | 0,39 |

TABLEAU IV.9 – Observation de la validation avec divers scénario de ventilation

On peut également observer les résultats en validation avec les divers scénarios (cf. tableau IV.9). Ainsi, on remarque que les meilleurs résultats sont toujours obtenus lorsque le scénario de ventilation est le même pour le jeu aillant servi à l'identification et pour le jeu ayant

servi à la validation. Cela paraît logique. On remarque néanmoins que les autres erreurs sont acceptables et que les erreurs les plus élevées apparaissent lorsque la différence de scénarios est vraiment importante entre l'identification et la validation (identification sur un scénario de ventilation et validation sur un scénario sans ventilation et vice versa).

L'ensemble de ces études montrent que de nombreux aspects peuvent venir impacter l'identification. En effet, dans certaines situations, une expertise devient nécessaire et le processus d'identification seul n'est pas suffisant. C'est pourquoi nous proposons un procédé de "méta-optimisation".

IV.4 Besoin de "méta-optimisation" et de tenir compte de l'incertitude sur les phénomènes

Il s'agit de piloter le processus d'optimisation de l'identification afin d'en améliorer les résultats. Pour cela, les études réalisées en amont sur les phénomènes non maîtrisés et les analyses sur les paramètres seront utilisées afin d'identifier le modèle des salles PREDIS MHI à l'aide de toutes les connaissances que nous avons.

IV.4.1 Amélioration de l'identifiabilité paramétrique

Nous avons pu voir précédemment que la dispersion et l'arrivée en butée des valeurs des paramètres identifiés étaient des aspects clés de l'identifiabilité. Nous nous sommes donc intéressés à l'étude de ces aspects dans la réduction de notre modèle.

IV.4.1-i Réduction de la dispersion paramétrique

Le fait de fixer les paramètres les moins sensibles n'ayant pas donné de résultats concluants, nous avons fait des essais en fixant les paramètres les plus dispersés. En effet, ces paramètres étant difficilement identifiables, ils peuvent rendre l'identification des autres paramètres plus difficile. Nous avons pu observer une réduction de la dispersion lorsque les paramètres les plus dispersés étaient fixés. Nous avons étudié le comportement de l'identification en fonction du nombre de paramètres à identifier.

Ayant à disposition plusieurs jeux de données expérimentaux collectés à divers moments de l'année, nous avons pu analyser la dispersion des paramètres en fonction du moment de l'année. En effet, nous avons vu que les comportements pouvaient changer en fonction de la saison, en est-il de même pour la dispersion ? Il est en effet pertinent d'étudier la dispersion des paramètres afin de savoir si ce sont toujours les mêmes qui sont dispersés ou non. Combien faut-il fixer de paramètres pour obtenir le meilleur résultat ? Ce nombre dépend-il du moment de l'année ? Comment automatiser ce procédé ?

Ce sont les questions que nous nous sommes posées et dont les réponses seront détaillées dans le chapitre 5.

IV.4.1-ii Analyse de l'arrivée en limite de validité des paramètres

Bien que l'aspect dispersif des paramètres traduise un réel problème d'identifiabilité, l'arrivée aux bornes du domaine de validité paramétrique est également un aspect à ne pas négliger dans l'identification.

Lors du choix de la structure d'un modèle et de ses paramètres, des bornes d'incertitudes peuvent être définies et viennent délimiter un domaine de validité où les valeurs ont une cohérence physique. Lorsque trop d'incertitudes sont présentes lors de la réalisation du modèle, on peut délimiter le domaine de recherche par des bornes larges afin de ne pas risquer de trop contraindre les valeurs par des bornes incertaines. Cependant, lorsque les valeurs des paramètres identifiées arrivent en butée, elles perdent leur cohérence physique.

Ces valeurs aberrantes peuvent venir perturber les résultats d'identification et les valeurs des autres paramètres du modèle. Un algorithme contraignant les paramètres arrivant en butée à leur valeur nominale à l'itération suivante a donc été testé. De même que précédemment se sont posées des questions à propos de l'amélioration du résultat, du temps de calcul et de la reproductibilité du résultat en fonction du nombre de paramètres à fixer et de la saison.

IV.4.2 Etude des phénomènes mal maîtrisés

L'analyse de l'influence des phénomènes non maîtrisés nous a également permis de quantifier l'influence des divers scénarios et perturbations anticipables ou non.

Ainsi, afin d'améliorer les résultats d'identification, nous avons utilisé diverses pondérations en fonction de :

- L'erreur obtenue à l'itération précédente ;
- Le moment (journée, nuit, week end).

Ceci permet de donner plus de poids aux moments qui sont maîtrisés.

Ainsi, l'identification des paramètres d'un modèle thermique peut s'avérer complexe et perturbée par de nombreux facteurs. Afin de mettre en place des scénarios de gestion optimale, une recalibration régulière du modèle peut-être proposée afin de représenter au mieux le comportement du bâtiment à un moment précis. Nous cherchons une méthode de recalibration robuste : le modèle doit être assez facilement et rapidement recalibré mais également proposer une valeur d'erreur acceptable en validation. Ceci permet d'anticiper le comportement du système bâtiment et d'en déduire des stratégies de gestion.

CHAPITRE V

Analyse de la dispersion paramétrique afin d'obtenir un modèle réduit permettant un recalibrage robuste

Table des matières

| | |
|--|-----|
| V.1 Description du principe | 107 |
| V.1.1 Algorithme | 107 |
| V.1.2 Scénario de mise en place | 108 |
| V.1.3 Besoins et pré-requis techniques | 110 |
| V.2 Mise en place de la méthode sur des jeux de données PREDIS MHI | 111 |
| V.2.1 Test de la méthode sur des jeux de données obtenus par simulation COMFIE Pléïade | 111 |
| V.2.2 Test de la méthode sur des jeux de données expérimentaux | 116 |
| V.3 Validation de la méthode avec un autre modèle de bâtiment | 126 |
| V.3.1 Description du modèle | 126 |
| V.3.2 Description du modèle physique : utilisation d'un réseau de résistances et capacités thermiques | 126 |
| V.3.3 Mise en place de la méthode d'analyse de la dispersion | 129 |
| V.4 Simulation de la méthode en scénario de gestion | 131 |
| V.5 Conclusions sur l'utilisation de la méthode de réduction par analyse de la dispersion paramétrique | 132 |
| V.5.1 Avantages et inconvénients de cette méthode | 132 |
| V.5.2 Conclusion sur la méthode et généralisation | 133 |

Dans le chapitre précédent, nous avons démontré les limites d'une calibration dite "classique". Un certain nombre d'estimations paramétriques sont réalisées et à l'issue de cette série d'estimations, ce sont les valeurs de paramètres estimées qui donnent l'erreur la plus faible qui sont conservées. Cette méthode donne des résultats avec une précision acceptable mais est longue en temps de calcul (puisqu'elle nécessite autant d'optimisations que d'estimations paramétriques).

La dispersion paramétrique et de l'erreur traduisent un problème d'identifiabilité. Une analyse de la dispersion a été menée et a permis une réduction de la complexité du modèle. Ceci peut améliorer la dispersion globale et les résultats d'estimation paramétrique. Cette méthode cherche à résoudre les problèmes d'identification du modèle en minimisant la dispersion. La dispersion globale vise à être ainsi réduite en fixant certains paramètres.

Cette méthode, qui explore l'espace de recherche et permet une meilleure compréhension du comportement du bâtiment, sera détaillée dans la première partie de ce chapitre. Ensuite, elle sera testée sur le modèle des salles PREDIS à l'aide de jeux de données simulées issus du logiciel de simulation thermique COMFIE Pléiade ainsi que sur des jeux expérimentaux issus des enregistrements capteurs. Enfin, cette méthode sera validée sur un autre modèle de bâtiment.

V.1 Description du principe

Lorsque l'on utilise une calibration classique, les résultats obtenus peuvent être très dispersés, comme nous avons pu le voir dans le chapitre III. L'estimation paramétrique réalisée avec une calibration classique (meilleur résultat parmi plusieurs estimations ayant des points de départ différents) n'est pas robuste. En effet, l'algorithme ne converge pas vers une erreur unique, ni vers des valeurs de paramètres identifiées uniques.

Il s'agit ici d'améliorer l'estimation paramétrique afin de rendre la recalibration plus robuste. Pour cela, nous cherchons dans un premier temps à réduire la dispersion. Cette dispersion traduit en partie un manque d'identifiabilité lié à la complexité du modèle par rapport à la quantité d'information présentes dans les signaux d'entrée (contenu spectral) mais également aux non-linéarités par rapport aux paramètres physiques, comme cela a été rappelé dans le chapitre IV.

Afin de réduire la complexité du modèle, les paramètres sont fixés par ordre de dispersion (du plus dispersé au moins dispersé) et les modèles réduits obtenus sont comparés afin de trouver le meilleur compromis pour la gestion anticipative (précision, robustesse, adaptabilité, ...).

V.1.1 Algorithme

Soit M_n , un modèle thermique à n paramètres. Le modèle est identifié m fois pour un jeu de données et donne m valeurs d'erreur, et m valeurs identifiées pour chaque paramètres.

Un indice de dispersion est calculé pour chaque paramètre selon l'équation suivante :

$$\forall i \in [1, \dots, n], \text{disp}_i = \frac{\sigma_i}{p_{\max}^i - p_{\min}^i} \quad (\text{V.1})$$

Avec p_{\max}^i et p_{\min}^i les valeurs des bornes limitant l'espace de recherche pour le paramètre i et σ_i l'écart type du paramètre i calculé selon l'équation suivante :

$$\forall i \in [1, \dots, n], \sigma_i = \sqrt{\frac{1}{m} \sum_{k=1}^m (x_i^k - \mu_i)^2} \quad (\text{V.2})$$

Où x_i^k représente la $k^{\text{ème}}$ valeur identifiée du paramètre i et μ_i sa valeur moyenne sur l'ensemble des valeurs identifiées.

Il s'agit de l'écart type des valeurs estimées normalisées. On peut ensuite comparer les indices de dispersion des paramètres pour divers jeux de données. Un classement moyen des paramètres est réalisé allant du paramètre le plus dispersé au moins dispersé. Les similitudes des classements pour différents jeux de données peuvent être observées.

Les paramètres sont ensuite fixés un par un et le modèle obtenu est testé : une série d'estimation paramétrique est réalisée. Nous cherchons à montrer que la réduction du modèle en fixant les paramètres les plus dispersés réduit la dispersion globale. En réalisant plusieurs optimisations pour chaque modèle réduit, nous observons l'amélioration du modèle en comparaison d'une estimation paramétrique de l'ensemble des paramètres. Nous cherchons à trouver le nombre de paramètres à fixer qui offre le meilleur compromis précision/robustesse compte tenu de l'application choisie. Dans le cadre de la gestion optimale, seront donc observés :

- le temps de calcul ;
- la précision en estimation et surtout en validation ;
- la dispersion de l'erreur (écart type pour plusieurs résultats d'erreurs obtenues entre mesure et simulation) ;
- la dispersion paramétrique (écart type pour plusieurs résultats d'estimation).

Si plusieurs jeux de données sont à disposition, l'estimation paramétrique sera réalisée et validée sur des jeux de données ayant un comportement de même saison (période avec ou sans chauffage), ce qui garantit a priori un contenu spectral similaire. Un jeu de données pourra être tronqué pour être utilisé en estimation et en validation.

V.1.2 Scénario de mise en place

L'analyse de la dispersion paramétrique est réalisée une fois et elle nous permet d'obtenir un modèle réduit (en terme de nombre de paramètres à identifier), comme nous pouvons l'observer sur le schéma de la figure V.1. Ce modèle réduit possède des paramètres fixes et des paramètres à identifier. Les paramètres à fixer peuvent varier en fonction des saisons et des situations (présence ou non d'occupants, scénarios de chauffage, météo, etc.). Une étude

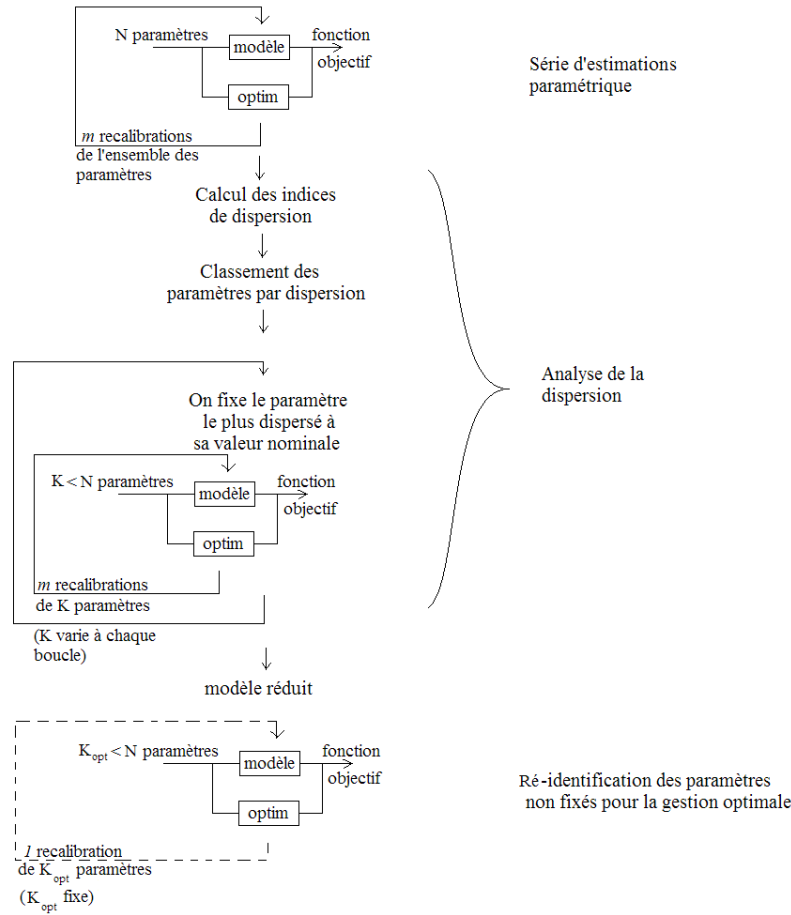


FIGURE V.1 – Mise en place de la réduction du modèle par analyse de la dispersion

préalable sur plusieurs jeux de données est à réaliser pour déterminer si la généralisation du modèle simplifié est possible ou non.

Le meilleur modèle réduit est déterminé selon les critères énoncés précédemment (temps de calcul, précision, dispersion). Une fois qu'il est obtenu, il est utilisé pour la gestion anticipative, c'est-à-dire qu'il sera fréquemment ré-identifié afin de représenter au mieux le comportement du bâtiment en fonction des situations et servira à prédire le comportement thermique du bâtiment à court horizon (24h, 48h ou 72h). En effet, cette analyse garantit un modèle léger (ayant une complexité réduite) et facilement ré-identifiable pour une précision adaptée à la gestion anticipative. De plus, si la dispersion est suffisamment réduite, alors une seule identification est suffisante pour arriver dans un intervalle d'erreur réduit. C'est pourquoi les m identifications du modèle réduit sont optionnelles (en pointillé sur la figure V.1).

Cette analyse requiert un expert afin de déterminer le modèle thermique retenu et de réaliser différents tests et observations. Une fois le nombre optimal de paramètres déterminé, si les paramètres à fixer varient d'un jeu de données à un autre, alors le modèle réduit peut-être re-calibré par un processus automatique qui calculera les nouveaux indices de dispersion pour une période donnée, comme on peut le voir sur le schéma de la figure V.2. Ce réajustement est plus rapide que l'analyse et plus simple à mettre en place, puisqu'une fois que le nombre

optimal de paramètres à identifier est connu, il suffit de réaliser une série d'identifications et de fixer les paramètres les plus dispersés.

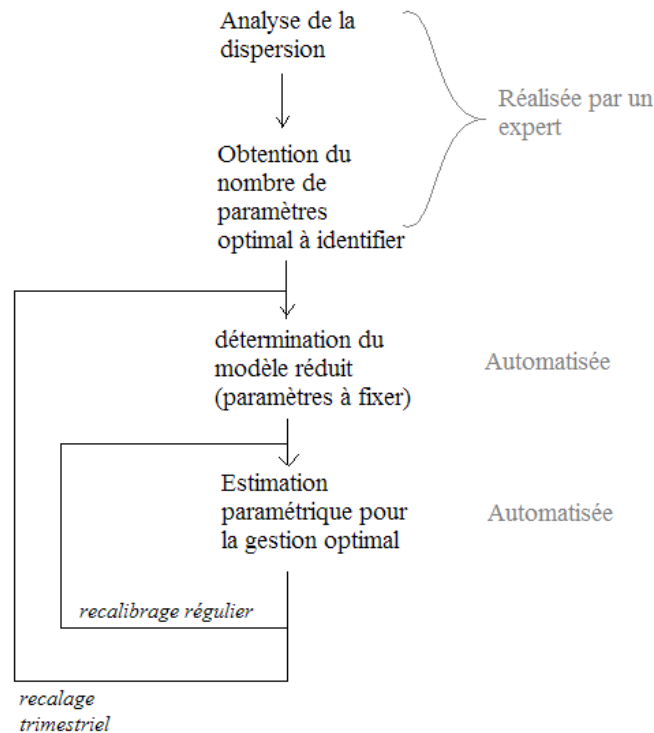


FIGURE V.2 – Mise en place de la réduction du modèle par analyse de la dispersion

V.1.3 Besoins et pré-requis techniques

Pour la mise en place de la gestion anticipative, un modèle thermique de bâtiment est nécessaire. Les valeurs initiales des paramètres du modèle sont calculées en fonction des caractéristiques du bâtiment.

Après obtention du modèle thermique, la mise en place de cette méthode nécessite quelques jeux de données afin de réaliser des calculs d'indices de dispersion et de pouvoir les comparer. Le bâtiment doit donc être équipé de capteurs qui collectent :

- températures (pièces adjacentes, températures intérieures et extérieure) ;
- flux solaires ;
- apports internes (puissance électrique, puissance de chauffage) ;
- concentration de CO_2 (pour évaluer l'apport de chaleur dû à l'occupation).

Lors de cette analyse, l'algorithme réalise m identifications par jeu de données puis évalue les indices de dispersion. Chaque modèle réduit (défini par un certain nombre de paramètres à identifier) est ensuite identifié m fois à nouveau afin de trouver le meilleur modèle réduit (en terme de nombre de paramètres à identifier). Soit n , le nombre de jeux de données à disposition, alors cette méthode d'optimisation requiert $2 \times n \times m$ optimisations. Le temps de calcul qui en découlent dépend de la longueur des jeux de données ainsi que du nombre de paramètres à identifier. Non seulement la réduction du nombre de paramètres à identifier

diminue le temps de calcul mais le fait de fixer les paramètres les plus difficiles à identifier permet également à l'optimisation de converger plus rapidement.

V.2 Mise en place de la méthode sur des jeux de données PREDIS MHI

Dans un premier temps, la méthode a été testée sur des jeux de données obtenus par simulation via le logiciel de simulation thermique COMFIE Pléïade. L'utilisation de ces jeux de données présentent plusieurs avantages :

- Ils sont disponibles sur de grandes plages de valeurs et à divers moments de l'année (été, hiver) ;
- Les scénarios de ventilation et d'occupation sont entièrement connus et définis ;
- Il n'y a pas d'erreur de capteurs (valeurs manquantes ou erronées).

Ceci permet de vérifier que la méthode fonctionne et que si des erreurs surviennent, elles ne sont pas liées à des problèmes de bruits ou de capteurs. Néanmoins, il conviendra de tester la méthode sur des jeux de données expérimentaux, et c'est ce qui sera réalisé dans un deuxième temps.

V.2.1 Test de la méthode sur des jeux de données obtenus par simulation COMFIE Pléïade

Il s'agit dans un premier temps de déterminer s'il y a un nombre optimal de paramètres à identifier qui peut convenir à des jeux de données de toute l'année. Ensuite, il est intéressant de regarder si les paramètres à fixer sont les mêmes en fonction des saisons ou non.

Comme dit précédemment, lorsque l'on utilise les jeux de données issus de simulation avec le logiciel COMFIE, les flux solaires ne sont pas pris en compte et nous avons donc un modèle à 15 paramètres et non pas 19.

V.2.1-i Étude préliminaire : Détermination du nombre optimal de paramètres à identifier

Une première expertise doit être réalisée sur plusieurs jeux de données afin de trouver le nombre de paramètres qui donnent le meilleur compromis pour une recalibration robuste.

Quatre jeux de données issus de saisons différentes ont été identifiés plusieurs fois. A la suite d'une analyse de la dispersion, les paramètres les plus dispersés sont fixés un par un puis les résultats obtenus pour chacun des modèles réduits sont comparés. Il s'agit de déterminer si un nombre optimal de paramètres à fixer et à identifier peut être trouvé et satisfaire les critères de la recalibration robuste :

- bonne identifiabilité : faible dispersion paramétrique ;

- erreur non dispersée : meilleur modèle garanti avec une seule identification, gain en temps de calcul ;
- précision de l'erreur de validation : afin d'avoir une bonne anticipation de la température intérieure.

Sur la figure V.3, on peut observer l'évolution de la dispersion de l'erreur en fonction du nombre de paramètres à identifier.

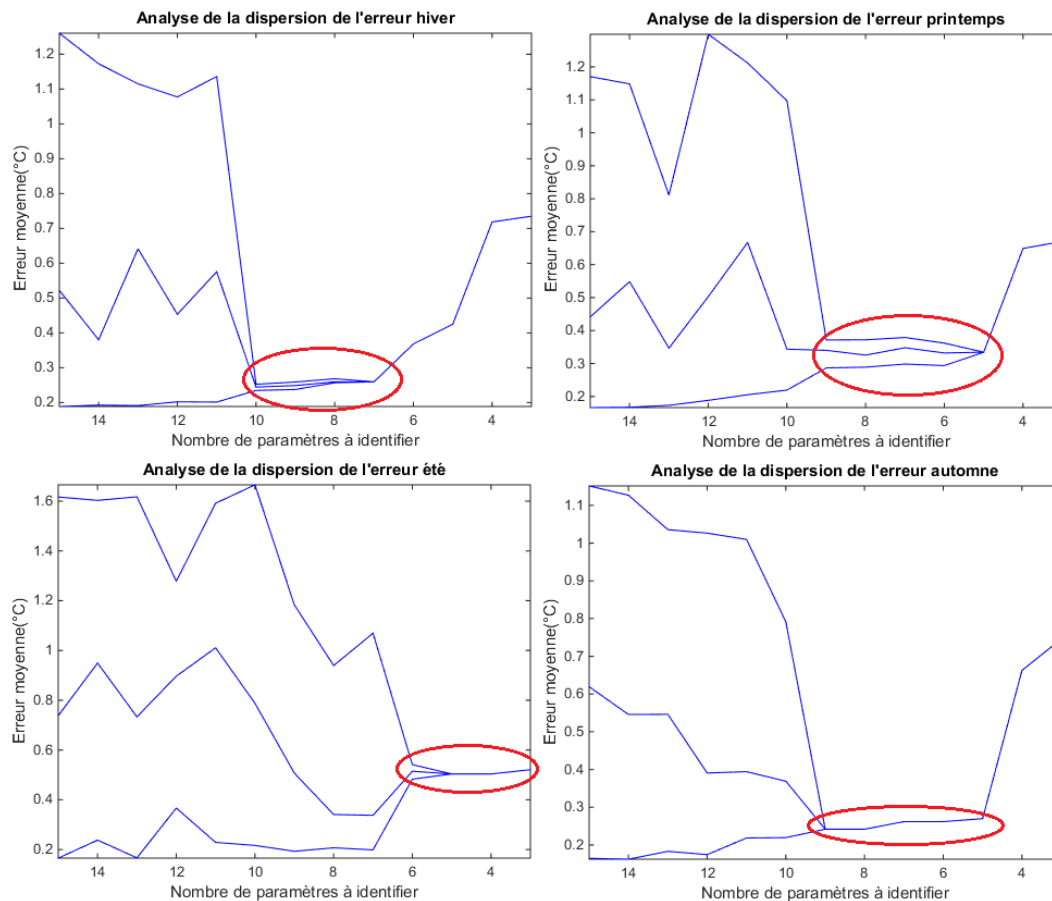


FIGURE V.3 – Observation de l'évolution de la dispersion de l'erreur pour chaque saison

Les cercles rouges montrent les modèles considérés comme acceptables. Le tableau V.1 montrent les modèles acceptables en terme de nombre de paramètres ajustables en fonction de la saison (pour les autres, les valeurs pré-déterminées seront utilisées).

| | Nombres de paramètres à identifier | | | | | | | | | | |
|-----------|------------------------------------|---|---|---|---|---|---|----|----|----|----|
| | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| hiver | | | | | | | | | | | |
| printemps | | | | | | | | | | | |
| été | | | | | | | | | | | |
| automne | | | | | | | | | | | |

TABLEAU V.1 – Modèles acceptables en terme de nombre de paramètres à identifier en fonction des saisons

Nous remarquons que lorsqu'il y a du chauffage, un nombre de paramètres plus important est nécessaire pour obtenir une bonne estimation paramétrique. Le faible nombre de

paramètres ajustables en été est probablement lié à la pauvreté spectrale des signaux (inoccupation et pas de chauffage). Un compromis pour l'ensemble des saisons peut toutefois être trouvé entre 6 et 7 paramètres ajustables puisque ces deux modèles sont acceptables pour trois des saisons sur quatre. Cependant, le modèle à 7 paramètres qui ne fait pas partie des modèles acceptables en été donne de moins bons résultats (en termes de dispersion) que le modèle à 6 paramètres de l'hiver.

L'évolution de la dispersion moyenne (somme des écarts types normalisées) des paramètres peut également être observée pour choisir le meilleur modèle.

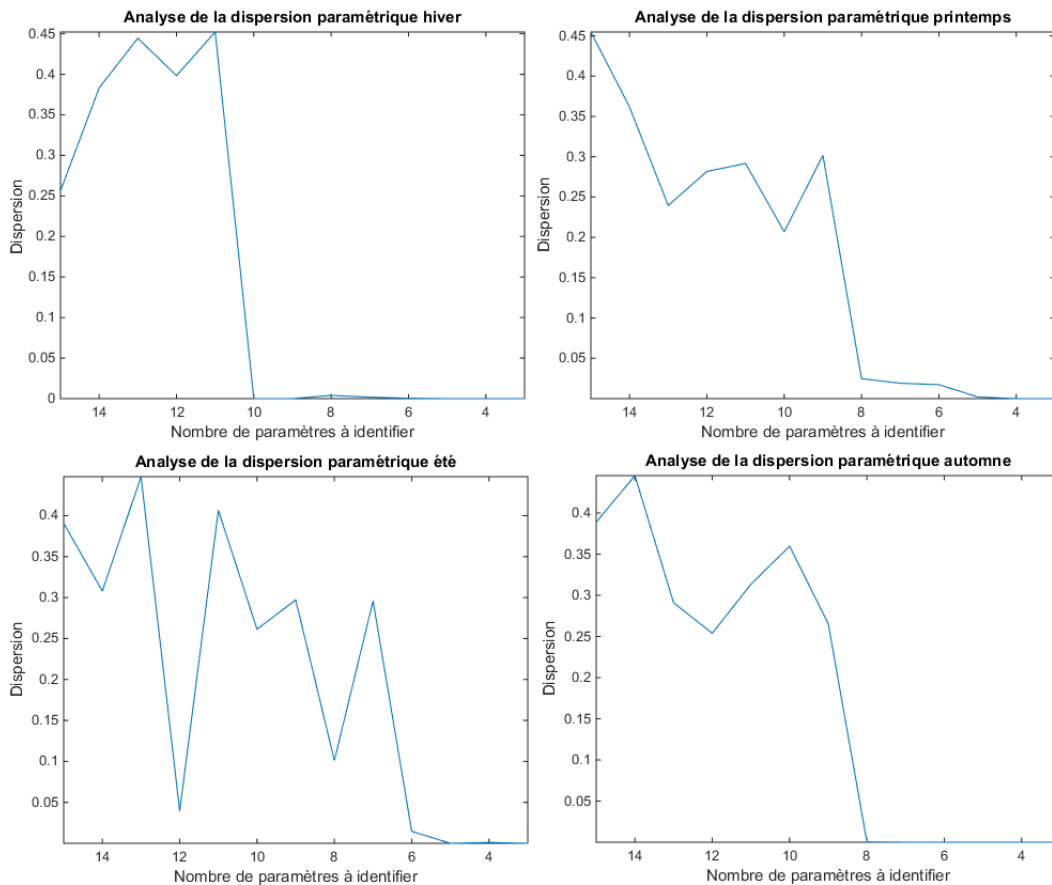


FIGURE V.4 – Observation de la dispersion paramétrique moyenne pour chaque saison

On remarque que, sur la figure V.4, la dispersion paramétrique moyenne est acceptable à partir de 10 paramètres ajustables pour le jeu d'hiver, 8 paramètres ajustables pour les jeux de printemps et d'automne, et 6 paramètres ajustables pour le jeu d'été.

Cette fois-ci, le modèle à 6 paramètres ajustables est le modèle qui offre le meilleur compromis. En effet, au dessus, certains jeux donnent des résultats trop dispersés, en dessous, nous savons qu'il y a une perte en précision (comme vu sur la figure V.3).

Voyons ce que donnent les résultats de validation pour les modèles à 6 et 7 paramètres à identifier. Les modèles réduits ont donc été identifiés sur des jeux de données et validés sur d'autres de même saison.

Le tableau V.2 contient les valeurs d'erreur moyenne et maximum atteintes en validation pour les quatre jeux de données pour les deux modèles réduits à 6 et à 7 paramètres. L'erreur

| | Modèle réduit à 6 paramètres | | Modèle réduit à 7 paramètres | |
|------------------|------------------------------|-----------------|------------------------------|-----------------|
| | erreur moyenne (°C) | erreur max (°C) | erreur moyenne (°C) | erreur max (°C) |
| Jeu d'hiver | 0,97 | 3,17 | 1,50 | 3,97 |
| Jeu de printemps | 0,29 | 0,98 | 0,30 | 0,94 |
| Jeu d'été | 0,55 | 1,57 | 0,32 | 1,88 |
| Jeu d'automne | 0,51 | 3,05 | 0,51 | 3,05 |

TABLEAU V.2 – Comparaison des résultats de validation en fonction des saisons pour les modèles réduits à 6 et 7 paramètres

moyenne est calculée en faisant la moyenne des différences en valeurs absolues entre mesure et simulation. L'erreur maximum correspond à l'écart le plus grand en valeur absolue. On remarque que les erreurs sont globalement proches. L'erreur moyenne du jeu d'hiver pour un modèle à 6 paramètres est 30 % plus faible que pour celui à 7 paramètres. Par contre, l'erreur du modèle à 6 paramètres pour le jeu d'été est moins bonne que celle du modèle à 7 paramètres, en revanche, l'erreur maximum est inférieure. Les résultats des jeux d'automne et de printemps sont égaux ou presque.

Ainsi, le modèle à 6 paramètres à identifier semble offrir le meilleur compromis pour l'ensemble des jeux de données.

V.2.1-ii Choix des paramètres à fixer

Maintenant que nous avons vu que le modèle ayant 6 paramètres à identifier offrait le meilleur compromis sur une année entière, voyons quels sont les paramètres fixés pour chaque jeu de données et s'il s'agit des mêmes.

Le tableau V.6 classe les paramètres par ordre de dispersion pour chaque jeu de données du plus dispersé (fixé en premier) au moins dispersé. La correspondance des paramètres par rapport aux numéros utilisés pour plus de clarté est référencée dans le tableau V.5.

| efficiency | coefVentil | Presidual | Roffices | R1 | Rcorridor | coefOccup | coefChauff | airLeak | C | RL | R0 | coefElec | Tw0 | Rspace |
|------------|------------|-----------|----------|----|-----------|-----------|------------|---------|---|----|----|----------|-----|--------|
| 8 | 15 | 10 | 2 | 6 | 3 | 12 | 14 | 11 | 7 | 4 | 5 | 13 | 9 | 1 |

FIGURE V.5 – Légende pour la figure de la dispersion paramétrique en fonction de la saison

| | du plus dispersé au moins dispersé | | | | | | | | | | | | | | |
|-----------|------------------------------------|----|----|----|----|----|----|----|----|----|----|----|----|---|---|
| hiver | 4 | 8 | 5 | 12 | 10 | 11 | 15 | 7 | 6 | 9 | 13 | 14 | 1 | 2 | 3 |
| printemps | 3 | 15 | 2 | 8 | 6 | 10 | 14 | 11 | 12 | 7 | 13 | 1 | 9 | 4 | 5 |
| été | 8 | 2 | 4 | 5 | 6 | 14 | 3 | 15 | 10 | 7 | 12 | 11 | 13 | 9 | 1 |
| automne | 8 | 2 | 15 | 6 | 7 | 10 | 14 | 3 | 12 | 11 | 13 | 9 | 1 | 5 | 4 |

FIGURE V.6 – Classement des paramètres par ordre de dispersion (du plus dispersé au moins dispersé) en fonction du moment de l'année

A droite de la figure V.6, on trouve les paramètres qui seront estimés et à gauche ceux qui seront fixés (puisque dispersés). On remarque des similitudes entre les modèles. En effet, les paramètres R_1 , $efficiency$, $P_{residual}$ et $coef_{Ventil}$ (6, 8, 10 et 15) font toujours partie des paramètres fixés quelque soit le moment de l'année. Il resterait donc à déterminer les cinq derniers paramètres à fixer pour obtenir le modèle réduit à six paramètres.

Pour cela, un recalage en terme de paramètres à fixer pour le modèle réduit pourrait être réalisé à intervalle régulier (une fois tous les 3 mois par exemple). Certains paramètres étant toujours très dispersés, ils pourraient être fixés tout le temps. Il ne resterait plus qu'à choisir les paramètres les plus dispersés parmi les paramètres restant pendant les phases de recalage du modèle réduit. Ainsi, 4 paramètres seraient fixés en permanence, et seulement 5 supplémentaires seraient à fixer afin d'obtenir un modèle à 6 paramètres à identifier. Ces 6 paramètres seraient choisis en fonction des moments de l'année (comme nous avons vu que les paramètres ajustables en fonction des saisons n'étaient pas tout à fait les mêmes).

V.2.1-iii Validation du modèle réduit sur un grand nombre de jeux de données issus de la simulation

La réduction du modèle à 6 paramètres à identifier a été validée sur un grand nombre de jeux de données simulés étalés sur l'année entière. Afin de juger de la pertinence de notre méthode, les résultats en précision et en temps de calcul seront comparés. Le tableau V.3 présente les résultats obtenus. Ainsi, un grand nombre de phases d'estimations paramétriques suivi d'une phase de validation (simulation avec les paramètres estimés) a été réalisé. Pour chaque saison, une phase d'analyse préalable pour déterminer les paramètres à fixer a également été mise en place.

Dans le tableau, les phases d'analyse (pour la deuxième méthode), d'identification et de validation sont distinguées. En bas du tableau, on peut voir les résultats globaux en erreur moyenne et en erreur maximum atteinte. Les deux méthodes comparées sont les suivantes :

- **Méthode de calibration classique** : l'ensemble des paramètres du modèle sont régulièrement ré-identifiés (la méthode d'optimisation consiste à conserver le jeu de paramètres qui donne l'erreur la plus faible parmi 10 estimations paramétriques) ;
- **Analyse de la dispersion paramétrique** : les paramètres du modèle réduit sont régulièrement ré-identifiés (avec une seule optimisation), le modèle réduit est recalé tous les 3 mois afin de déterminer les paramètres à fixer.

| | | Méthode de calibration classique (19 paramètres ajustables) | Estimation paramétrique avec analyse de la dispersion préalable (6 paramètres ajustables) |
|-------------------------|-----------------------|--|--|
| Phase d'analyse | erreur moyenne | x | 0,3 |
| | erreur maximum | x | 2,6 |
| phases d'identification | erreur moyenne | 0,5 | 0,6 |
| | erreur maximum | 11,3 | 2,2 |
| phases de validation | erreur moyenne | 1,0 | 0,6 |
| | erreur maximum | 11,3 | 2,9 |
| Total | erreur moyenne | 0,7 | 0,5 |
| | erreur maximum | 11,3 | 2,9 |

TABLEAU V.3

En utilisant le modèle réduit issu de l'analyse de la dispersion, l'erreur est globalement plus faible, notamment durant les phases de validation, ce qui est ce que nous recherchons afin de garantir une calibration robuste. De plus, l'erreur maximum atteinte est relativement faible

(moins de 3°C) avec le modèle réduit alors qu'elle atteint des valeurs extrêmes (supérieure à 11°C) dans le cas de la calibration classique.

De plus, cette analyse préliminaire de la dispersion paramétrique permet de ré-estimer rapidement les paramètres du modèle réduit du fait que le modèle converge plus rapidement et, comme nous l'avons vu précédemment, la dispersion de l'erreur en est réduite également, ce qui permet de ne réaliser qu'une seule optimisation pour trouver le jeu de paramètres qui minimise l'erreur. Ainsi, le temps de simulation globale de la calibration classique pour l'ensemble des jeux de données disponibles (une année de simulation) est d'environ trois heures et demi tandis que le temps de simulation du modèle réduit avec réajustement régulier des paramètres à fixer est d'une quarantaine de minutes seulement. Ceci s'explique par le fait que la calibration classique nécessite plusieurs estimations à chaque optimisation tandis que dans le cas où le modèle réduit est utilisé pour la ré-estimation des paramètres du modèles, une seule estimation est réalisée. De plus, le modèle réduit converge plus rapidement, puisque les éléments qui perturbaient l'identification (les paramètres dispersés) ont été fixés.

Ainsi, ces premiers résultats sont très encourageants quant à la qualité de la méthode. Voyons maintenant si la méthode donne de bons résultats pour des jeux de données expérimentaux. Ces jeux de données sont obtenus par des capteurs disposés dans les salles PREDIS MHI.

V.2.2 Test de la méthode sur des jeux de données expérimentaux

Nous disposons de huit jeux de données expérimentaux collectés sur la plateforme PREDIS MHI :

- Quatre jeux de données avec du chauffage (jeux d'hiver) :
 - février 2014 (13 jours) ;
 - décembre 2014 (11 jours) ;
 - janvier 2014 (28 jours) ;
 - février 2015 (8 jours).
- Quatre jeux de données sans chauffage (jeux d'été) et sans clim (il n'y a pas de climatisation dans la salle de classe) :
 - avril 2014 (19 jours) ;
 - juin 2014 (11 jours) ;
 - juillet 2014 (8 jours) ;
 - août 2014 (8 jours).

Ici, le nombre d'optimisations est fixé à $m = 10$ et le modèle possède $n = 19$ paramètres. Dix optimisations sur le modèle avec l'ensemble des paramètres ont été réalisées pour chacun des 8 jeux de données (les jeux de données servent également à la validation lorsque l'identification a été réalisée sur un autre jeu de données). A partir d'une série d'optimisations, on peut calculer les indices de dispersions des paramètres. Ceux-ci sont observés sur les figures [V.7](#) et [V.8](#).

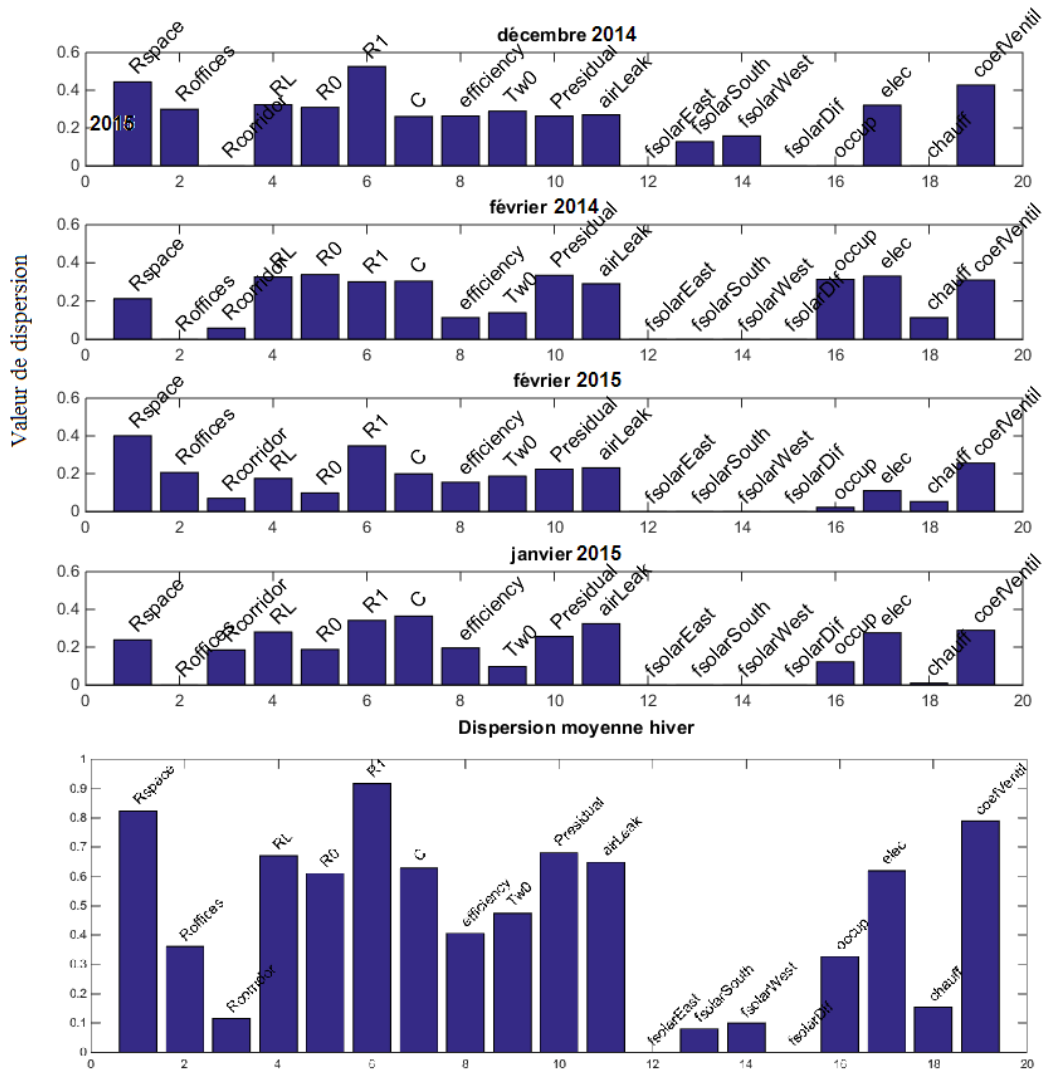


FIGURE V.7 – Dispersion des paramètres pour les jeux de données d’hiver de PREDIS et dispersion moyenne hiver

On remarque des similitudes globales :

- les paramètres R_{space} , R_1 et $coefVentil$ sont souvent parmi les plus dispersés ;
- les coefficients de flux solaire ne sont quasiment jamais dispersés.

Mais également des similitudes par saison :

- Le coefficient multiplicateur de l’apport de chaleur dû au flux électrique est dispersé en hiver et non dispersé en été : c’est un apport qui est négligeable par rapport au chauffage et à l’occupation en hiver, alors que ce n’est pas le cas en été ;
- Le coefficient multiplicateur de l’apport de chaleur dû à l’occupation est dispersé en été et peu dispersé en hiver : il fait en effet partie des apports prépondérants en hiver, de plus, les salles ne sont pas toujours occupées en été (juillet et août) ;

C’est pourquoi l’étude a ici été menée séparément sur les jeux d’été et les jeux d’hiver. Puisque les profils sont proches par saison, la moyenne des dispersions normalisées a été

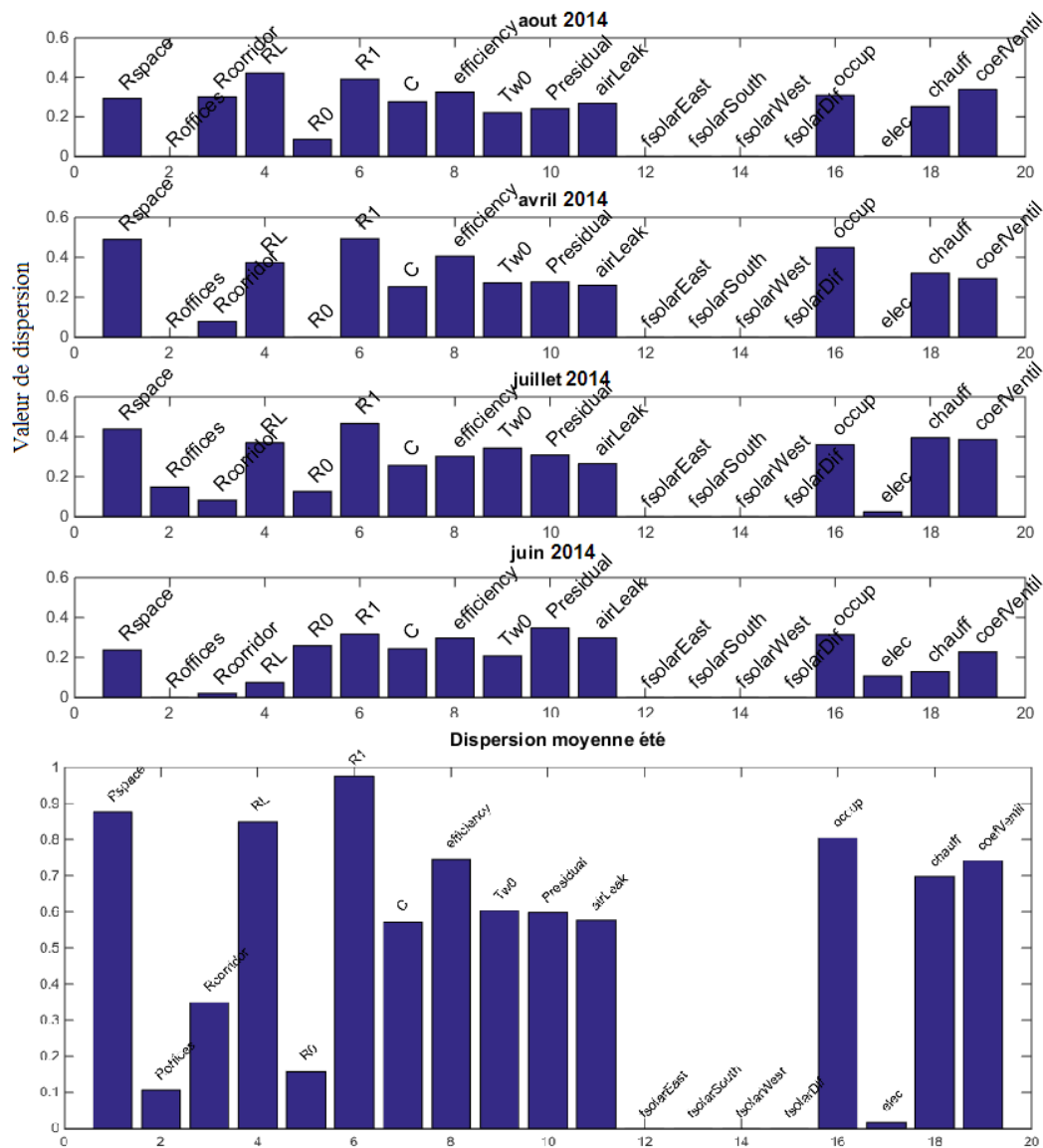


FIGURE V.8 – Dispersion des paramètres pour les jeux de données d’été de PREDIS et dispersion moyenne été

réalisée pour déterminer l’ordre dans lequel seront fixés les paramètres. Pour chaque jeu de données, en fonction de la saison, les paramètres sont fixés un par un dans l’ordre de dispersion visible dans le tableau V.4. Les modèles réduits obtenus sont identifiés dix fois. Les modèles possèdent donc de moins en moins de degrés de liberté.

Dans le tableau V.4, les paramètres en rouge sont ceux dont l’indice de dispersion est très différent en fonction de la saison, les autres paramètres (en noir), ont des indices de dispersion similaire.

Pour chaque modèle, nous nous intéresserons :

- au comportement de l’erreur :
 - dispersion de l’ensemble des résultats obtenus ;
 - erreur moyenne du meilleur modèle ;
 - erreur maximum du meilleur modèle.

| Hiver | Été |
|---------------------|---------------------|
| R_1 | R_1 |
| R_{space} | R_{space} |
| $coef_{Ventil}$ | R_L |
| $P_{residual}$ | $coef_{occup}$ |
| R_L | $efficiency$ |
| air_{Leak} | $coef_{Ventil}$ |
| C | $coef_{chauff}$ |
| $coef_{elec}$ | T_{w0} |
| R_0 | $P_{residual}$ |
| T_{w0} | air_{Leak} |
| $efficiency$ | C |
| $R_{offices}$ | $R_{corridor}$ |
| $coef_{occup}$ | R_0 |
| $R_{corridor}$ | $R_{offices}$ |
| $coef_{chauff}$ | $coef_{elec}$ |
| f_{solar}^{West} | f_{solar}^{East} |
| f_{solar}^{South} | f_{solar}^{South} |
| f_{solar}^{East} | f_{solar}^{West} |
| f_{solar}^{Diff} | f_{solar}^{Diff} |

TABLEAU V.4 – Ordre de dispersion moyenne des paramètres par saison, des plus dispersés (en haut) au moins dispersés (en bas)

- au comportement des paramètres : dispersion des valeurs identifiées pour chaque optimisation.
- à l'erreur obtenue en validation sur un autre jeu de donnée.

V.2.2-i Analyse des jeux d'hiver

V.2.2.i - a) Analyse du comportement de l'erreur d'identification

On peut observer l'évolution de la dispersion de l'erreur en fonction du nombre de paramètres à identifier sur la figure V.9.

La dispersion de l'erreur de chaque modèle est observable verticalement, les courbes minimum et maximum délimitant les valeurs des erreurs obtenues au cours des 10 identifications. La courbe du milieu correspond à la moyenne des erreurs. Ainsi, plus l'écart est grand et plus le modèle est dispersé. Les modèles à 12 et 11 paramètres ajustables semblent offrir les meilleurs compromis entre précision et faible dispersion pour l'ensemble des jeux de données. En effet, la dispersion est faible et l'erreur moyenne obtenue est faible, ce sont des modèles acceptables.

Cependant, cette valeur correspond à une erreur moyenne sur la durée de simulation. Voyons si l'erreur reste en dessous d'un seuil acceptable à tous les instants. La figure V.10 rend compte des histogrammes des erreurs obtenues entre température simulée et température mesurée en fonction du nombre de paramètres à identifier pour le jeu de données de Décembre 2014. On peut observer que l'erreur des modèles à 12 et 11 paramètres est contenue dans un intervalle plus petit et atteint un maximum à 1,5°C tandis que le modèle à 19 paramètres donne des valeurs d'erreurs allant jusqu'à 2,5°C.

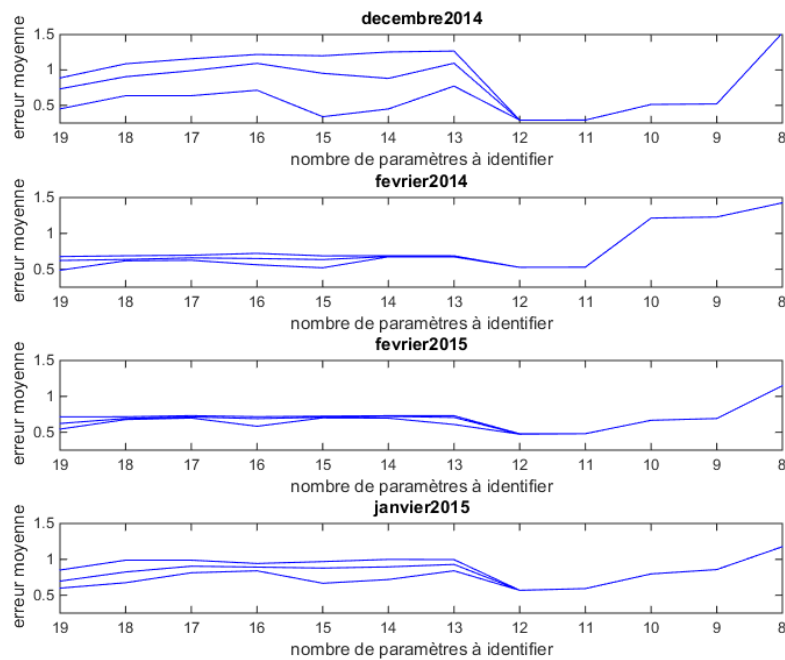


FIGURE V.9 – Évolution de la dispersion de l'erreur pour les jeux d'hiver de PREDIS

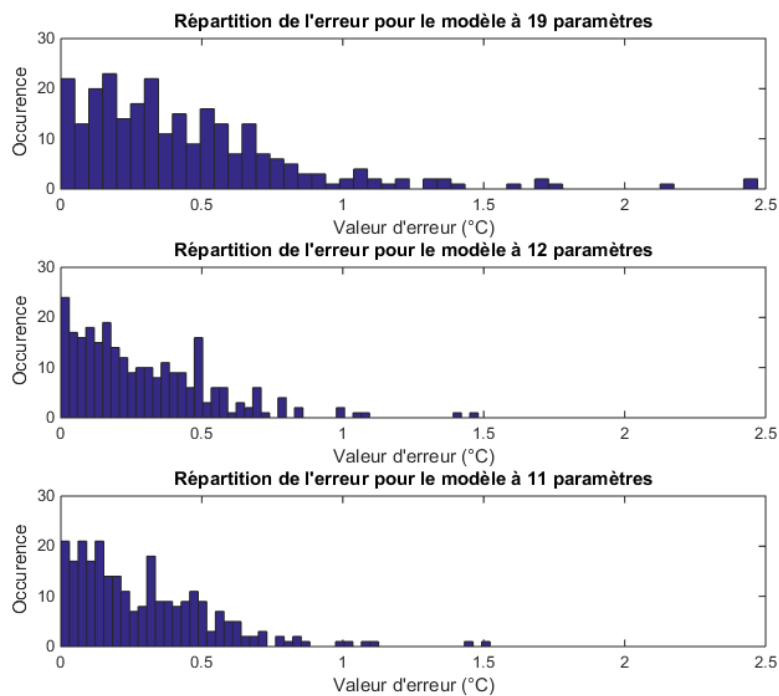


FIGURE V.10 – Évolution de l'histogramme de l'erreur en fonction du nombre de paramètres à identifier pour le jeu de décembre 2014

La réduction du nombre de paramètres améliore donc l'erreur moyenne mais également l'erreur maximum dans ce cas.

V.2.2.i - b) Analyse du comportement des paramètres

Voyons comment évolue la dispersion moyenne des paramètres en fonction du nombre de paramètres à identifier. La dispersion paramétrique globale est obtenue en faisant la moyenne

des écarts types normalisés en fonction du domaine de recherche (cela correspond à la moyenne des indices de dispersion).

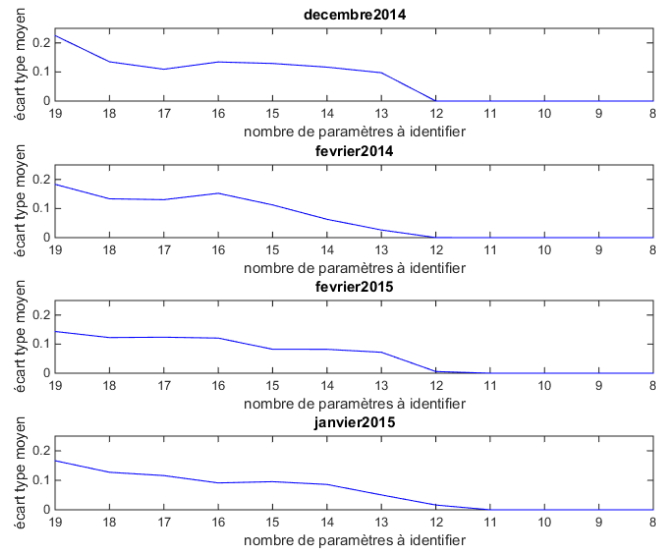


FIGURE V.11 – Évolution de la dispersion paramétrique moyenne pour les jeux d’hiver de PREDIS

De même que précédemment, à partir du modèle ayant 12 paramètres à identifier, on obtient une dispersion quasiment nulle (voir figure V.11).

V.2.2.i - c) Observation des résultats en validation

Les résultats d’identification des modèles à 19 et à 12 paramètres obtenus pour un jeu de données ont été simulés pour les trois autres jeux de données d’hiver. Les résultats de validation croisée sont présentés dans le tableau à deux entrées V.5. Le tableau qualifie la validation du modèle à 12 paramètres par rapport au modèle à 19 paramètres.

| | | Validation | | | |
|----------------|---------------|--------------|---------------|--------------|--------------|
| | | Février 2014 | Décembre 2014 | Janvier 2015 | Février 2015 |
| Identification | Février 2014 | x | meilleure | équivalente | équivalente |
| | Décembre 2014 | meilleur | x | meilleure | meilleure |
| | Janvier 2015 | meilleure | meilleure | x | meilleure |
| | Février 2015 | moins bonne | moins bonne | moins bonne | x |

TABEAU V.5 – Comparaison des résultats en validation du modèle à 12 paramètres par rapport au modèle à 19 paramètres

Globalement, les résultats sont meilleurs ou équivalents, sauf pour le jeu de février 2015. Ceci peut s’expliquer par le fait que le jeu de données de février 2015 est relativement court (8 jours) car comprenant une période de vacances. Ce jeu contient donc des informations assez pauvres par rapport aux autres jeux de données qui sont plus long et qui contiennent des situations différentes dans un même jeu (périodes de vacances et de non vacances). Or nous avons remarqué que la réduction du modèle à 12 paramètres à identifier améliorerait les résultats d’identification (précision et dispersion de l’erreur). Ainsi, simuler un jeu de données avec des valeurs de paramètres identifiées issus du modèle réduit d’une situation très

spécifique peut donner de moins bon résultats que le modèle avec tous les paramètres, qui reste plus globale dans l'estimation paramétrique.

C'est d'ailleurs ce qu'on observe sur les courbes de la figure V.12 qui montrent les résultats obtenus en validation sur le jeu de décembre 2014 pour une identification réalisée avec le jeu de février 2015. Les 7 derniers jours de ce jeu de données (décembre) correspondent aux vacances scolaires. Ici, nous remarquons que la validation du modèle à 19 paramètres traduit globalement correctement la tendance de la courbe de température, tandis que le modèle à 12 paramètres est très mauvais en prédiction sur les 5 premiers jours, mais devient très bon ensuite, pendant la période de vacances.

Ainsi, en réduisant le modèle sur ce jeu de données qui correspond à un comportement très précis, on donne plus de poids à ce comportement, et l'estimation paramétrique donnera de moins bon résultats en validation.

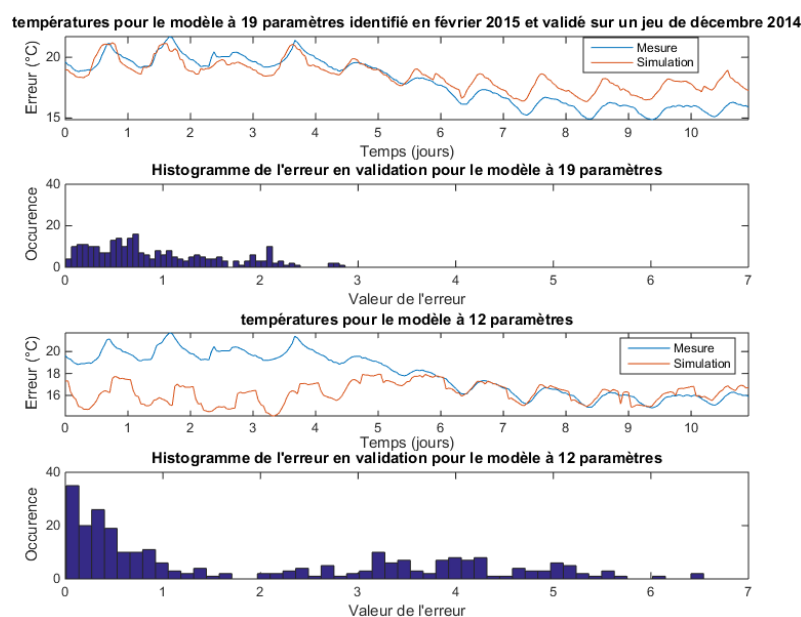


FIGURE V.12 – Comparaison des modèles à 19 et à 12 paramètres identifiés sur un jeu de février 2015 et simulés sur un jeu de décembre 2014

Le reste des résultats de prédiction est globalement meilleur, comme on peut le voir sur les courbes de la figure V.13. Ici, l'identification a été réalisée sur le jeu de décembre 2014 pour une validation avec le jeu de janvier 2015.

V.2.2-ii Analyse des jeux d'été

De même que précédemment, le modèle est réduit en fixant un par un les paramètres du plus dispersé au moins dispersé.

V.2.2.ii - a) Analyse du comportement de l'erreur d'identification

L'évolution de la dispersion du critère à minimiser pour les quatre jeux de données est observable sur la figure V.14.

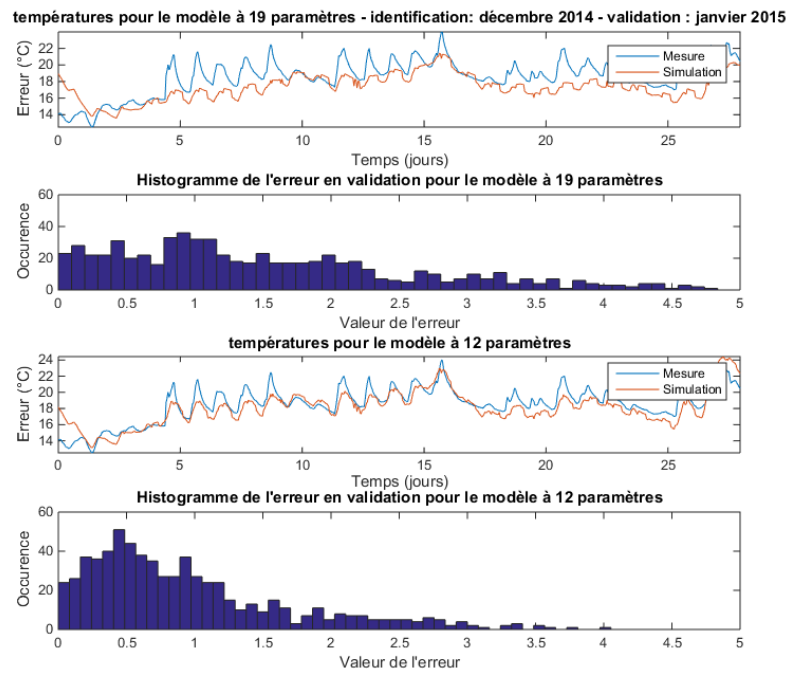


FIGURE V.13 – Comparaison des modèles à 19 et à 12 paramètres identifiés sur un jeu de décembre 2014 et simulés sur un jeu de janvier 2015

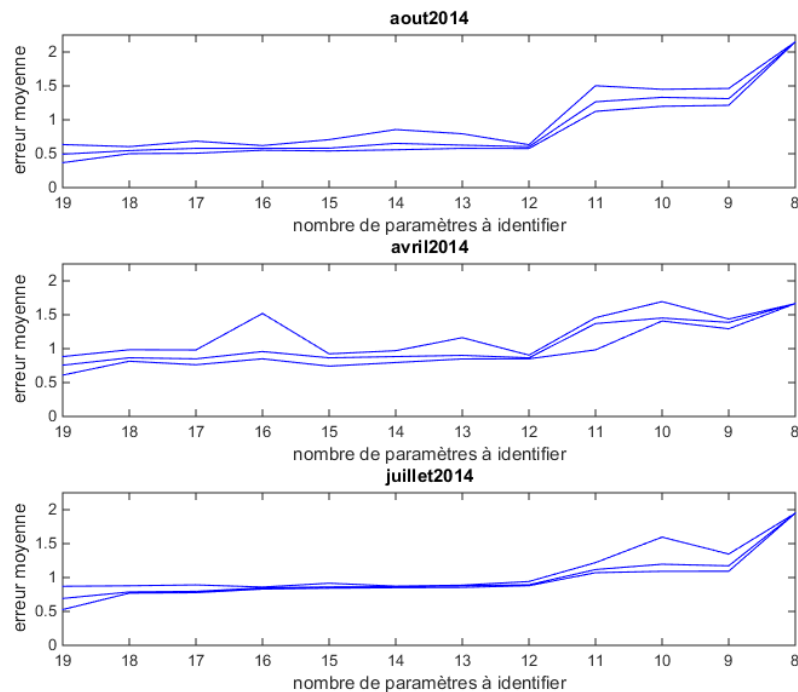


FIGURE V.14 – Évolution du critère de dispersion pour les jeux d'été de PREDIS

L'amélioration du modèle n'est pas aussi flagrante qu'avec les jeux de données hiver. Toutefois, le modèle à 12 paramètres à identifier pourrait convenir.

Si on compare les histogrammes d'erreur des modèles à 19, 12 et 11 paramètres sur la figure V.15 on voit qu'à partir de 11 paramètres à identifier, l'erreur est vraiment dégradée. Pour le modèle à 12 paramètres, la précision est moins bonne mais reste acceptable.

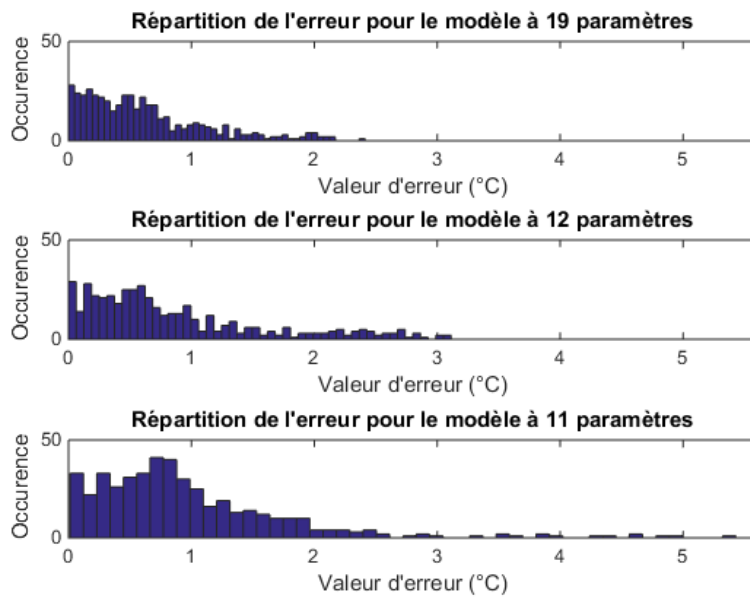


FIGURE V.15 – Évolution de l'histogramme de l'erreur en fonction du nombre de paramètres à identifier pour le jeu d'avril 2014

V.2.2.ii - b) Analyse du comportement des paramètres

L'évolution de la dispersion paramétrique (figure V.16) semble légèrement décroître. Cependant, au delà de 12 paramètres à identifier, la dispersion paramétrique augmente à nouveau.

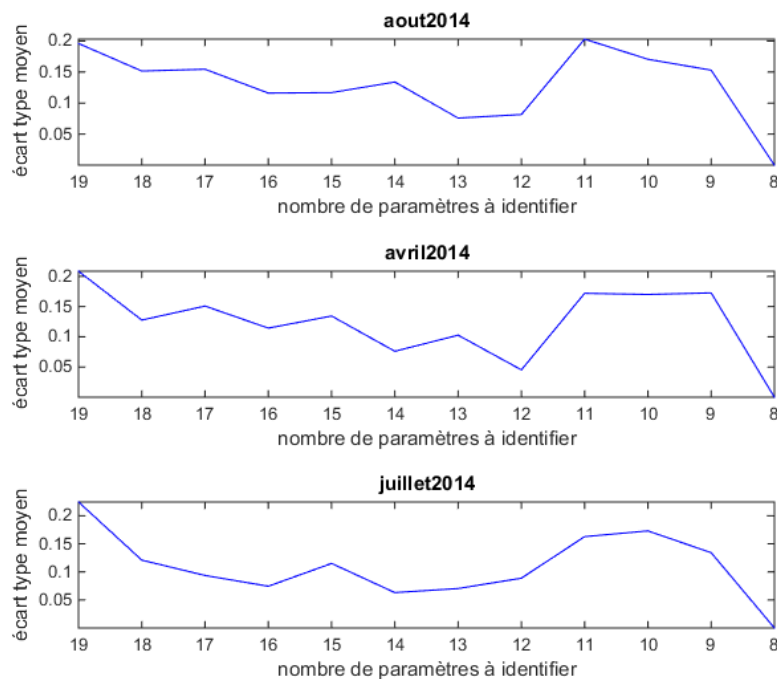


FIGURE V.16 – Évolution de la dispersion paramétrique moyenne pour les jeux d'été de PREDIS

V.2.2.ii - c) Observation des résultats en validation

Comparons maintenant les résultats en prédiction pour les modèles à 19 et 12 paramètres à identifier. Sur la figure V.17 qui montre plusieurs résultats de validation et la comparaison des modèles à 19 et à 12 paramètres, on remarque que les erreurs sont légèrement moins bonnes mais restent acceptables.

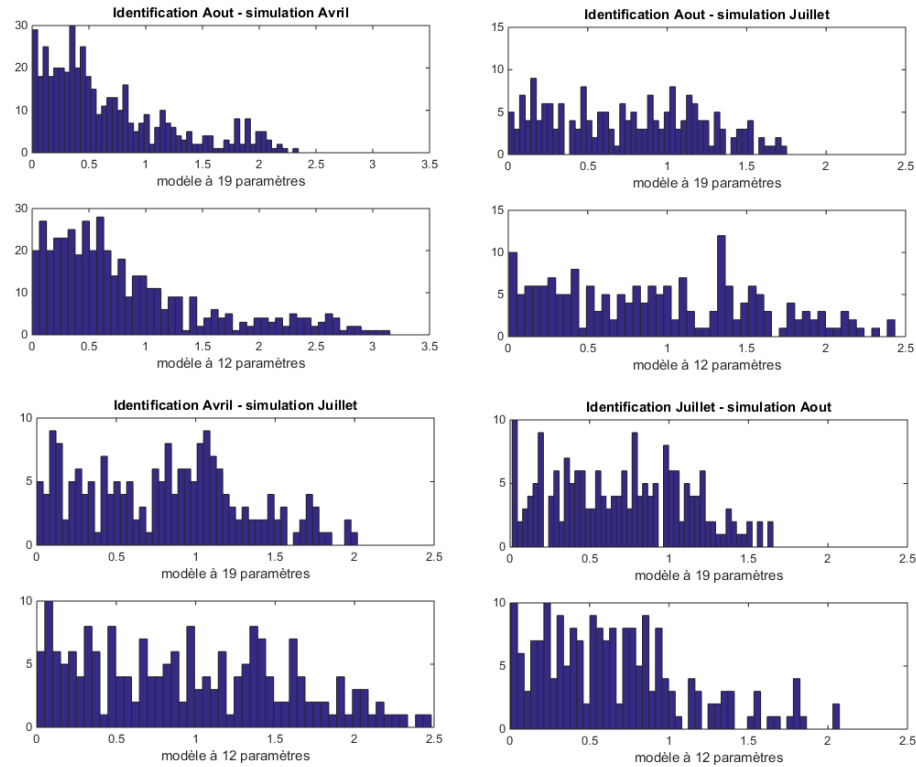


FIGURE V.17 – Histogramme d’erreur pour les modèles à 19 et 12 paramètres en validation pour divers jeux d’été de PREDIS

Ainsi, malgré la légère perte de précision, le modèle à 12 paramètres offre malgré tout un modèle valide.

V.2.2-iii Conclusions

Ainsi, les résultats de la méthode sur les jeux de données expérimentaux sont plus mitigés qu’avec les jeux de données simulés mais néanmoins relativement bons. Les mauvais résultats obtenus en validation ont pu être expliqués par le contenu spectral très spécifique d’un jeu de données et également par la différence importante du nombre de jours utilisés pour l’estimation et pour la validation. Dans le cadre de la gestion anticipative, l’anticipation de la température intérieure se fait sur du court terme (24 à 48h) et utilise des jeux données contigus à la période de validation, ce qui limite ce genre de situations spécifiques.

Afin d’être validé, la méthode doit être testé sur d’autres modèles de bâtiments. C’est ce qui a été fait avec un modèle de bureau situé dans les bâtiments G-SCOP à Grenoble. Il s’agit cette fois-ci d’un bureau plus petit avec un mur en contact avec l’extérieur, ce qui n’était pas le cas des salles PREDIS MHI. Il est ainsi très intéressant de voir si la méthode fonctionne aussi pour ce genre de situations.

V.3 Validation de la méthode avec un autre modèle de bâtiment

V.3.1 Description du modèle

L'espace étudié qui a servi à la validation de la méthode est un bureau situé dans le laboratoire G-SCOP à Grenoble. Il s'agit d'un bureau contenant 4 tables de travail. L'occupation moyenne est d'environ 3 personnes et varie en fonction des périodes de l'année.

Cette pièce possède un mur en contact avec l'extérieur qui comprend une fenêtre et un appareil de chauffage, deux murs en contact avec le couloir dont un qui comprend une porte ainsi qu'un mur en contact avec les escaliers qui comprend un placard et une issue de secours. L'un des murs en contact avec le couloir dispose d'une fenêtre. Le sol et le plafond sont en contact avec des bureaux ou des salles de classes.

Ainsi, à la différence des salles PREDIS, cette pièce est directement en contact avec l'extérieur, d'où l'intérêt pour la validation de notre étude.

Le schéma [V.18](#) montre un plan de coupe vu du haut du bureau.

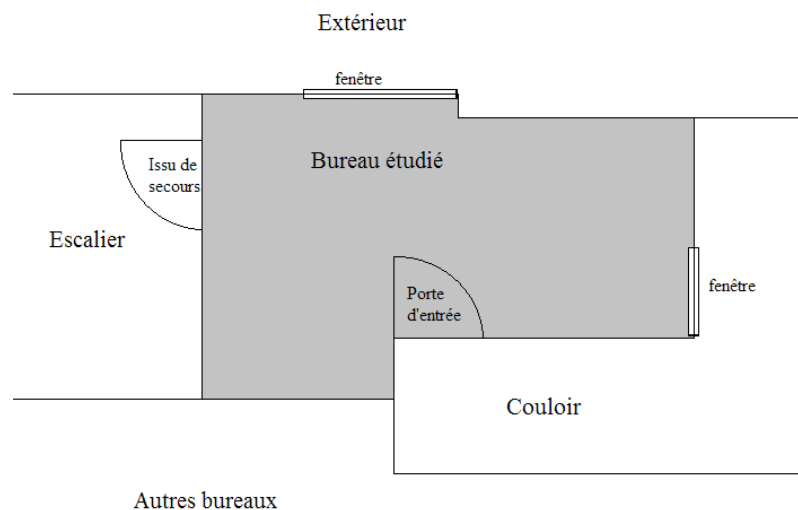


FIGURE V.18 – Plan en coupe vu du dessus du bureau étudié

V.3.2 Description du modèle physique : utilisation d'un réseau de résistances et capacités thermiques

Le modèle aéraulique est visible sur la figure [V.19](#).

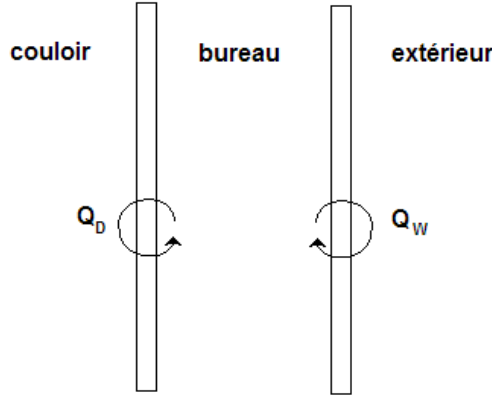


FIGURE V.19 – Schéma aéraulique du bureau

Ceci permet d'obtenir les équations d'échange d'air du bureau avec l'extérieur et le couloir :

$$\begin{aligned}
 Q_W &= Q_W^0 + \zeta_W Q_W^1 \\
 Q_D &= Q_D^0 + \zeta_D Q_D^1 \\
 R_D &= \frac{1}{\rho_{air} c_{p,air} Q_D} \\
 R_W &= \frac{1}{\rho_{air} c_{p,air} Q_W} \\
 T_n - T_{in} &= R_n \varphi_n \\
 T_{in} - T_{out} &= R_{out} \varphi_{out} \\
 \zeta_D (T_n - T_{in}) &= R_D \varphi_D \\
 \zeta_W (T_{in} - T_{out}) &= R_W \varphi_W \\
 T_{in} - \tau &= R_i \varphi_i \\
 C_i \frac{d\tau}{dt} &= \varphi_i \\
 \varphi_n + \varphi_D + \varphi_{in} &= \varphi_i + \varphi_{out} + \varphi_W
 \end{aligned} \tag{V.3}$$

avec $c_{p,air} = 1004$, $\rho_{air} = 1.204$, $\zeta_W = 0$ lorsque la fenêtre est fermée et $\zeta_D = 0$ lorsque la porte est fermée.

Ceci conduit aux équations de modèle suivantes :

— **Le modèle thermique :**

$$\text{Soit } \frac{1}{R} = \frac{1}{R_i} + \frac{1}{R_{out}} + \frac{\zeta_W}{R_W} + \frac{1}{R_n} + \frac{\zeta_D}{R_D}$$

$$\begin{aligned}
 \frac{d\tau}{dt} &= \frac{R - R_i}{R_i^2 C_i} \tau + \frac{R}{R_i C_i} \left(\frac{1}{R_{out}} + \frac{\zeta_W}{R_W} \right) T_{out} + \frac{R}{R_i C_i} \left(\frac{1}{R_n} + \frac{\zeta_D}{R_D} \right) T_n \\
 T_{in} &= \frac{R}{R_i} \tau + R \varphi_{in} + R \left(\frac{1}{R_{out}} + \frac{\zeta_W}{R_W} \right) T_{out} + R \left(\frac{1}{R_n} + \frac{\zeta_D}{R_D} \right) T_n
 \end{aligned} \tag{V.4}$$

avec R_n , R_{out} , R_i et C_i invariants dans le temps.

— **Le modèle aéraulique :**

$$C_i \frac{d\Gamma_{in}}{dt} = -(Q_W + Q_D)\Gamma_{in} + Q_W\Gamma_{out} + Q_D\Gamma_n + S_{CO_2}^{body}n \quad (V.5)$$

où Γ_{in} , Γ_{out} et Γ_n représentent les concentrations en CO_2 du bureau, de l'extérieur et du couloir.

La modélisation thermique de cette pièce a été réalisée à l'aide d'un circuit de résistances et de capacités thermiques, comme on peut le voir sur le schéma V.20.

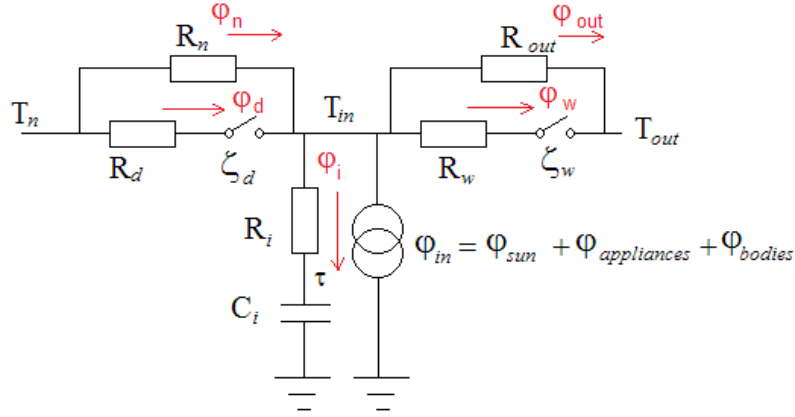


FIGURE V.20 – Schéma électrique équivalent pour le bureau

Le contact de la pièce avec les autres bureaux et couloir est représenté par les branches de gauche du schéma. La branche avec un interrupteur correspond à la porte du bureau qui peut être ouverte ou fermée et donc favoriser les échanges avec les zones adjacentes. De même, la branche de droite représente le mur en contact avec l'extérieur, et la branche avec l'interrupteur correspond à l'état d'ouverture de la fenêtre sur l'extérieur.

Le tableau V.6 décrit les paramètres du modèle à identifier. Il y en a 10.

| | |
|------------------|---|
| C_i | Capacité de l'air |
| Q_{out}^0 | Débit d'air initial vers l'extérieur |
| $Q_{corridor}^0$ | Débit d'air initial vers le couloir |
| Q_W | Débit d'air de la fenêtre |
| Q_D | Débit d'air de la porte |
| R_{out} | Résistance du mur vers l'extérieur |
| R_n | Résistance du mur vers le couloir |
| $Solar_{Factor}$ | Coefficient multiplicateur du flux solaire |
| R_i | Résistance de l'air |
| K_{heat} | Coefficient multiplicateur du flux de chauffage |

TABLEAU V.6 – Description des paramètres du modèle physique du bureau

Les variables ζ_d et ζ_w sont obtenues à l'aide de capteurs, de même que les températures.

V.3.3 Mise en place de la méthode d'analyse de la dispersion

V.3.3-i Première étape : réduction du nombre de paramètres du modèle par analyse de la dispersion

Il s'agit de vérifier que la réduction du modèle à identifier dans le cadre de la gestion anticipative améliore la robustesse de la recalibration et que nous obtenons de bons résultats en validation.

Le protocole de test suivra donc le même schéma que précédemment :

- Analyse de la dispersion globale sur plusieurs jeux de données ;
- Choix du modèle réduit offrant le meilleur compromis ;
- Validation du modèle réduit sur d'autre jeu de donnée.

V.3.3.i - a) Étude préliminaire : analyse de la dispersion paramétrique

Il s'agit de trouver le nombre de paramètres à identifier qui améliore l'identification. La figure V.21 montre les résultats concernant l'évolution de l'erreur pour trois jeux de données différents issus de données expérimentales du bureau étudié.

On remarque que les modèles réduits qui pourraient convenir à tous les jeux de données ont 7 ou 8 paramètres.

Le tableau V.7 rend compte de l'ordre dans lequel les paramètres sont fixés.

| Jeu de données 1 | Jeu de données 2 | Jeu de données 3 |
|------------------|------------------|------------------|
| $Solar_{Factor}$ | $Q0_{corridor}$ | $Q0_{corridor}$ |
| K_{heat} | C_i | K_{heat} |
| C_i | $Solar_{Factor}$ | C_i |
| $Q0_{corridor}$ | K_{heat} | $Solar_{Factor}$ |
| QD | QD | QD |
| R_i | R_n | R_n |
| R_n | R_i | R_i |
| $Q0_{out}$ | $Q0_{out}$ | $Q0_{out}$ |
| QW | QW | QW |
| R_{out} | R_{out} | R_{out} |

TABLEAU V.7 – Classement des paramètres par ordre de dispersion, du plus dispersé au moins dispersé, pour trois jeux de données différents

On remarque que les 4 paramètres les plus dispersés sont toujours les mêmes : $Solar_{Factor}$, K_{heat} , C_i et $Q0_{corridor}$. Les modèles réduits à 7 et 8 paramètres auront donc 2 ou 3 paramètres fixés parmi ces quatre là.

Voyons maintenant si ces modèles réduits donnent de bons résultats en validation.

V.3.3.i - b) Validation des modèles réduits

Afin de juger de la robustesse de la calibration des modèles, les modèles à 10 et à 7 paramètres à identifier ont été comparés. A l'issue d'une identification sur deux semaines, la validation sera réalisée sur un jeu de deux semaines.

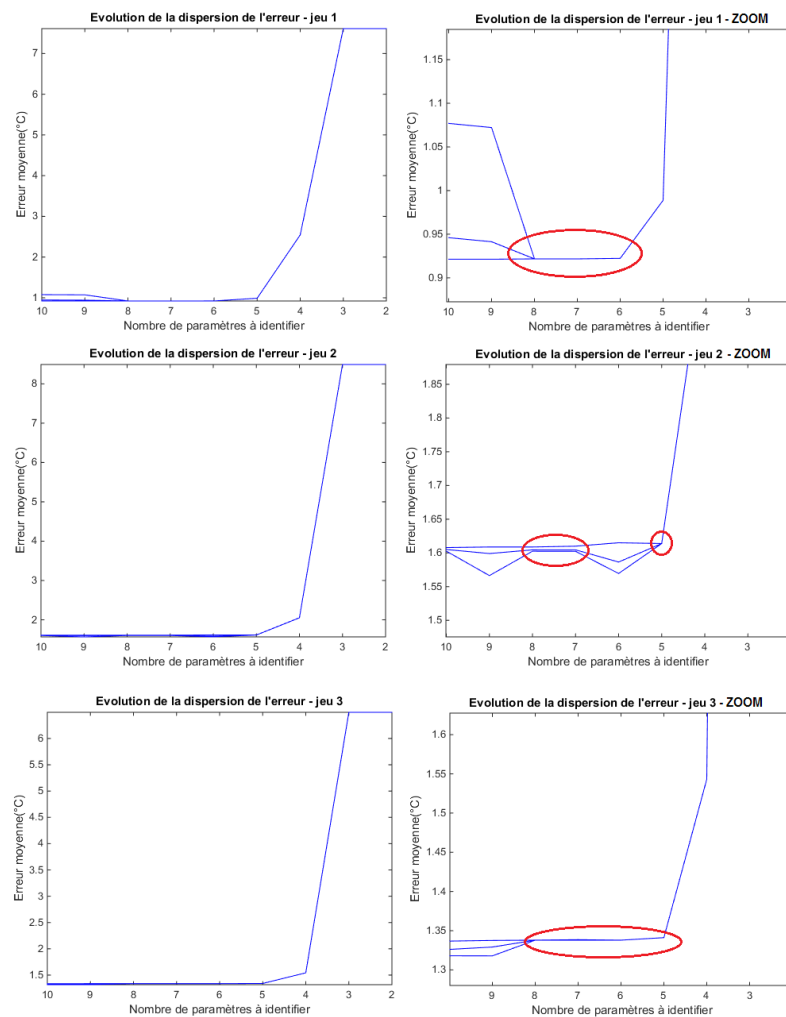


FIGURE V.21 – Évolution de la dispersion de l'erreur en fonction du nombre de paramètres à identifier

Les résultats obtenus sont présentés dans le tableau V.8.

| | | Modèle avec les 10 paramètres | Modèle réduit à 7 paramètres |
|-------|----------------|-------------------------------|------------------------------|
| Jeu 1 | Erreur moyenne | 0,86 | 0,85 |
| | Erreur max | 3,5 | 3,5 |
| Jeu 2 | Erreur moyenne | 0,99 | 0,98 |
| | Erreur max | 4,14 | 4,13 |
| Jeu 3 | Erreur moyenne | 0,86 | 0,87 |
| | Erreur max | 6,32 | 6,03 |

TABLEAU V.8 – Résultat en validation du modèle avec l'ensemble des paramètres et du modèle réduit à 7 paramètres

Comme nous pouvons le voir, les résultats en validations sur les trois jeux de données sont très proches.

En effet, une différence de trois paramètres fixés ne présente pas d'amélioration significative quant à la calibration. Cependant, nous avons vu précédemment que les modèles réduits étaient moins dispersés que le modèle initial à 10 paramètres. Ainsi, pour des résultats en validation équivalents, la recalibration peut être réalisée avec un modèle réduit qui sera plus

rapide à converger. En fonction des jeux de données, les paramètres fixés sont $Solar_{Factor}$, K_{heat} et C_i ou $Q_{corridor}^0$, C_i et $Solar_{Factor}$ ou $Q_{corridor}^0$, K_{heat} et C_i .

V.4 Simulation de la méthode en scénario de gestion

Lorsque l'on veut réaliser de la gestion anticipative, on souhaite anticiper la température intérieure à court terme afin de mettre en place les scénarios de gestion (ventilation, chauffage) optimaux.

Ici, nous avons choisi d'anticiper la température à deux jours en utilisant les 7 jours précédents pour l'estimation paramétrique. Ainsi, le recalibrage du modèle s'effectue tous les deux jours. Pour cela, les données d'un an simulées avec le logiciel COMFIE Pléiades ont été utilisées.

Deux scénarios ont été comparés afin de justifier de l'utilité de l'analyse de la dispersion paramétrique :

- **Scénario standard** : c'est le modèle initial (avec l'ensemble des paramètres à utiliser) qui est utilisé et régulièrement recalibré (*la meilleure estimation parmi 10* tous les 7 jours) pour anticiper la température des 2 prochains jours ;
- **Scénario utilisant la méthode de réduction** : c'est le modèle réduit qui est régulièrement recalibré (*une estimation paramétrique* tous les 7 jours) pour anticiper la température des 2 prochains jours. Ce modèle réduit est régulièrement recalé (tous les 3 mois), c'est-à-dire qu'on refait une analyse de dispersion pour fixer les paramètres les plus dispersés.

Ces deux scénarios sont comparés en terme de rapidité et de précision en validation puisque c'est le recalibrage robuste qui est le critère le plus important dans le cadre de la gestion anticipative.

Le tableau V.9 montre les valeurs d'erreurs moyennes et maximum obtenues pour différentes phases de la régulation pour les deux scénarios.

Scénario standard

| | Erreur moyenne | Erreur maximum |
|-------------------------|----------------|----------------|
| Phases d'identification | 0,44 | 6,23 |
| Phases de validation | 0,63 | 8,38 |

Scénario utilisant la méthode de réduction

| | Erreur moyenne | Erreur maximum |
|-------------------------|----------------|----------------|
| Phases d'analyse | 0,32 | 2,6 |
| Phases d'identification | 0,42 | 2,4 |
| Phases de validation | 0,51 | 3,63 |

TABLEAU V.9 – Erreurs moyennes et maximum obtenues pour différentes phases de la gestion pour les deux scénario

L'erreur moyenne ainsi que l'erreur maximum atteinte sont plus faibles pour le scénario avec la réduction par analyse de la dispersion. Ceci est une garantie intéressante quant à la robustesse du modèle pour la gestion anticipative.

De plus, comme nous pouvons le voir dans le tableau V.10, la simulation globale sur l'année est bien plus faible dans le cas d'un scénario avec analyse de dispersion (il y a un rapport de 18 entre les deux temps).

| | Scénario standard | Scénario avec analyse de dispersion |
|---------------------|-------------------|-------------------------------------|
| Temps de simulation | 93 275 secondes | 4888 secondes |

TABLEAU V.10 – Temps total de simulation sur une année pour les deux différents scénarios de gestion

Ceci garantit un modèle léger qui est fiable en terme de précision (erreur moyenne), de robustesse (erreur maximum) et qui est rapide à recalibrer.

V.5 Conclusions sur l'utilisation de la méthode de réduction par analyse de la dispersion paramétrique

V.5.1 Avantages et inconvénients de cette méthode

Le but de cette méthode est de trouver un modèle réduit qui offre le meilleur compromis pour représenter toutes sortes de situations et à divers moments de l'année. Ceci présente des avantages et des inconvénients.

V.5.1-i Inconvénients

- Le temps de calcul de l'expertise (détermination du nombre optimal de paramètres et meilleurs paramètres à conserver pour l'identification) est conséquent étant donné qu'il nécessite une série d'identification pour tester chaque modèle réduit.
- On peut avoir des erreurs en validation lorsque le jeu de données servant à l'estimation paramétrique correspond à une situation trop particulière. Pour corriger cela, on pourra ponctuellement agir sur l'estimation en conservant l'ensemble des paramètres pour l'identification ou utiliser des valeurs identifiées précédemment pendant une période plus riche en informations.
- Cette étude se focalise sur la dispersion paramétrique et non sur la cohérence des valeurs estimées. Ainsi, les paramètres sont considérés comme identifiables si leur valeurs identifiées ne sont pas dispersées sur l'espace de recherche, mais les valeurs peuvent être inconsistantes. C'est sur cet aspect que se focalisera la méthode de méta-optimisation qui sera détaillée dans la partie suivante.
- Une étude préalable de la dispersion des paramètres est nécessaire, et cela requiert plusieurs jeux de données en amont. A l'instar de la calibration classique.

V.5.1-ii Avantages de l'analyse de dispersion

- L'analyse de la dispersion et la réduction du modèle permet de réduire la dispersion globale. Ceci rend le modèle plus robuste à la ré-identification.
- Une fois le modèle réduit déterminé, le temps de calcul de la recalibration paramétrique est très court puisqu'il y a moins de paramètres à identifier.
- De plus, puisque la dispersion est réduite, alors une seule estimation paramétrique est suffisante pour trouver le meilleur modèle, ce qui réduit d'autant plus le temps de calcul lors des étapes d'estimation paramétrique.
- Une fois que le nombre optimal de paramètres est déterminé pour une année, l'automatisation du réajustement des paramètres à fixer est facile à mettre en place.

V.5.2 Conclusion sur la méthode et généralisation

A travers plusieurs tests sur divers jeux de données, nous avons pu remarquer qu'il était possible de trouver un nombre de paramètres optimal à conserver pour l'estimation paramétrique dans le cadre d'une gestion anticipative. Cependant, nous avons remarqué que les paramètres à fixer n'étaient pas les mêmes en hiver et en été. Pour automatiser le choix des paramètres à conserver, on peut classer les paramètres par indice de dispersion sur étude préalable d'un jeu de donnée et fixer les paramètres les plus dispersés à leur valeur nominale. Le modèle réduit ainsi obtenu servira à la gestion anticipative et sera fréquemment recalibré.

Les limites de cette méthode résident dans la difficulté de modéliser des sollicitations spécifiques lorsqu'on a identifié les paramètres sur une période courte où cette situation n'intervient pas. Confronté à ce genre de situation, on pourra pondérer l'estimation paramétrique en fonction des résultats précédemment obtenus et également en fonction des périodes ayant des scénarios connus (vacances scolaires par exemple).

Néanmoins, étant moins dispersé que le modèle global avec tous les paramètres, ce modèle permet un gain en temps de calcul non négligeable puisqu'une seule identification est suffisante pour trouver le modèle optimal. En effet, l'erreur obtenue pour plusieurs identifications est contenue dans un intervalle restreint, ce qui n'est pas le cas du modèle ayant tous les paramètres à identifier.

Enfin, la cohérence physique des valeurs n'est pas étudiée avec cette méthode, c'est pourquoi nous nous sommes intéressés à cette notion dans le prochain chapitre.

CHAPITRE VI

Pilotage optimal de l'estimation paramétrique par méta-optimisation (systématisation)

Table des matières

| | |
|--|-----|
| VI.1 Présentation de la méta-optimisation | 136 |
| VI.1.1 Intérêt de l'utilisation de la méta-optimisation | 136 |
| VI.1.2 Description du principe | 137 |
| VI.1.3 Algorithme | 139 |
| VI.1.4 Pré-requis | 140 |
| VI.2 Mise en place de la méthode sur des jeux de données PREDIS MHI | 141 |
| VI.2.1 Simulation sur un an de jeux de données simulés issus du logiciel COMFIE | 141 |
| VI.2.2 Validation de la méta-optimisation sur les jeux de données expérimentaux des salles PREDIS MHI | 144 |
| VI.3 Validation de la méthode avec un autre modèle de bâtiment | 149 |
| VI.4 Simulation de la méthode en scénario de gestion | 149 |
| VI.5 Conclusion | 151 |
| VI.6 Couplages des deux méthodes | 152 |
| VI.6.1 Principe | 152 |
| Conclusion générale | 158 |

Dans ce chapitre, nous nous sommes focalisés sur la cohérence des valeurs estimées des paramètres. En effet, lorsque les domaines des paramètres sont bornés par des valeurs limites de l'acceptabilité physique et que l'estimation paramétrique mène à une valeur arrivant aux bornes d'un domaine, comment peut-on justifier physiquement la valeur obtenue ? Cela pose la question de la validité physique du modèle obtenu. Nous avons cherché à guider l'algorithme d'optimisation afin d'éviter de glisser dans des zones où la fonction objectif est monotone du moins dans les directions paramétriques. La validité physique peut conduire à une plus grande capacité d'extrapolation du modèle à d'autres situations, car les zones de l'espace de recherche trop éloignées de la réalité pourrait donner ponctuellement de bons résultats d'identification mais être mauvais en prédiction. Nous proposons une méthode de méta-optimisation qui pilote les optimisations de l'estimation paramétrique. Ce procédé s'adapte à tout type d'algorithme d'optimisation.

Cette méthode exploite la connaissance physique des valeurs paramétriques connues a priori. Elle se focalise sur l'étude de la monotonie du critère à minimiser afin de détecter les paramètres ne permettant pas d'atteindre un minimum et d'orienter l'optimisation.

VI.1 Présentation de la méta-optimisation

Contrairement à la méthode décrite précédemment, il ne s'agit pas d'obtenir un modèle réduit plus facilement recalibrable mais de piloter l'optimisation afin d'obtenir de meilleurs résultats et cela, à chaque estimation des paramètres du modèle. Nous pouvons qualifier cette méthode de méta-optimisation, un procédé qui va piloter l'optimisation afin de l'orienter dans l'espace de recherche et ainsi éviter les zones où l'on estime que le critère n'admet pas de minimum et que les paramètres sont hors des zones physiquement acceptables.

Cette méthode sera dans un premier temps décrite et observée sur plusieurs jeux de données issus de simulations et de mesures expérimentales. Il s'agira également de vérifier que la prédiction à l'aide de la méta-optimisation donne de meilleurs résultats qu'à l'issue d'une calibration dite classique (meilleurs résultats d'identification parmi n). A l'aide de jeux de données ayant des scénarios différents (ventilation, occupation, chauffage), la capacité d'adaptabilité du modèle en fonction de la méthode d'estimation paramétrique utilisée sera évaluée.

VI.1.1 Intérêt de l'utilisation de la méta-optimisation

Le but de cette méthode est d'obtenir un recalibrage rapide tout en conservant une bonne précision en prédiction.

La conservation de la cohérence physique du modèle permet également d'augmenter la capacité d'extrapolation à d'autres situations. En effet, lorsque le modèle possède beaucoup de degrés de liberté, alors l'optimisation qui minimisera l'écart entre mesures et simulations peut entraîner l'estimation paramétrique vers des valeurs éloignées de la réalité. Ainsi, bien que l'estimation donne de très bons résultats en identification, la précision en prédiction ne sera pas garantie pour autant.

VI.1.2 Description du principe

Cette méthode de méta-optimisation vise à rejeter les paramètres sortant des zones physiquement acceptables. En effet, les bornes du domaine de recherche sont calculées à partir des valeurs nominales que l'on multiplie et divise par des valeurs qui dépendent de l'incertitude que nous avons sur les calculs. Ces valeurs sont subjectives et déterminées par un expert. Ces bornes de recherche larges permettent de laisser une grande liberté à l'algorithme pour trouver des valeurs qui minimisent l'écart entre mesure et simulation. Cependant, les valeurs qui arrivent en butée du domaine de recherche n'ont plus de sens physique.

Dans un premier temps, le modèle est évalué pour les valeurs nominales (calculées analytiquement). Ensuite, l'algorithme réalise une série d'itérations. Chaque itération est constituée de deux étapes :

- une analyse qui observe la monotonie du critère en fonction de chaque paramètre et juge de l'ajustabilité du paramètre ;
- une optimisation qui réalise une estimation paramétrique des paramètres ajustables.

Voici quelques définitions utiles :

Ajustable : Un paramètre est dit ajustable lorsque la courbe d'erreur en fonction de ce paramètre présente un minimum dans le domaine de validité où sont cherchés les paramètres (le fait d'être ajustable ou non est donc également lié au choix des bornes).

Valeur courante : La valeur courante d'un paramètre est égale à la dernière valeur estimée par l'algorithme et non arrivée hors du domaine acceptable. Si le paramètre sort du domaine acceptable alors sa valeur courante devient sa valeur a priori qui correspond à la valeur estimée à un pas précédent où le paramètre était ajustable. Si le paramètre n'a jamais été ajustable, il est rapporté à sa valeur nominale, évalué par l'expert.

VI.1.2-i Analyse : étude de la monotonie du critère de dispersion

Afin de juger de l'ajustabilité d'un paramètre, nous étudions la monotonie du critère à minimiser. La monotonie du critère dépend de la variation du paramètre observé mais également des valeurs des autres paramètres et donc du point de l'espace de recherche. La monotonie du critère en fonction d'un paramètre traduit la capacité de ce paramètre à minimiser le critère et donc son identifiabilité. En effet, si le critère est strictement monotone, alors la valeur de paramètre qui le minimisera sera vraisemblablement hors du domaine acceptable. Dans ce cas, le paramètre est considéré comme non ajustable en ce point de l'espace. Ainsi, le critère peut être monotone par rapport à un paramètre en un point de l'espace donnée, et ne le sera plus en un autre point. L'algorithme cherche à piloter l'optimisation afin de trouver le meilleur compromis entre précision et cohérence physique.

La figure [VI.1](#) montre l'évolution du critère à minimiser en fonction des variations d'un paramètre pour divers points de l'espace de recherche. Chaque couleur de courbe correspond à un point de l'espace de recherche où les variations du critère sont observées pour un paramètre qui varie tandis que tous les autres sont fixés. Ainsi on remarque qu'en fonc-

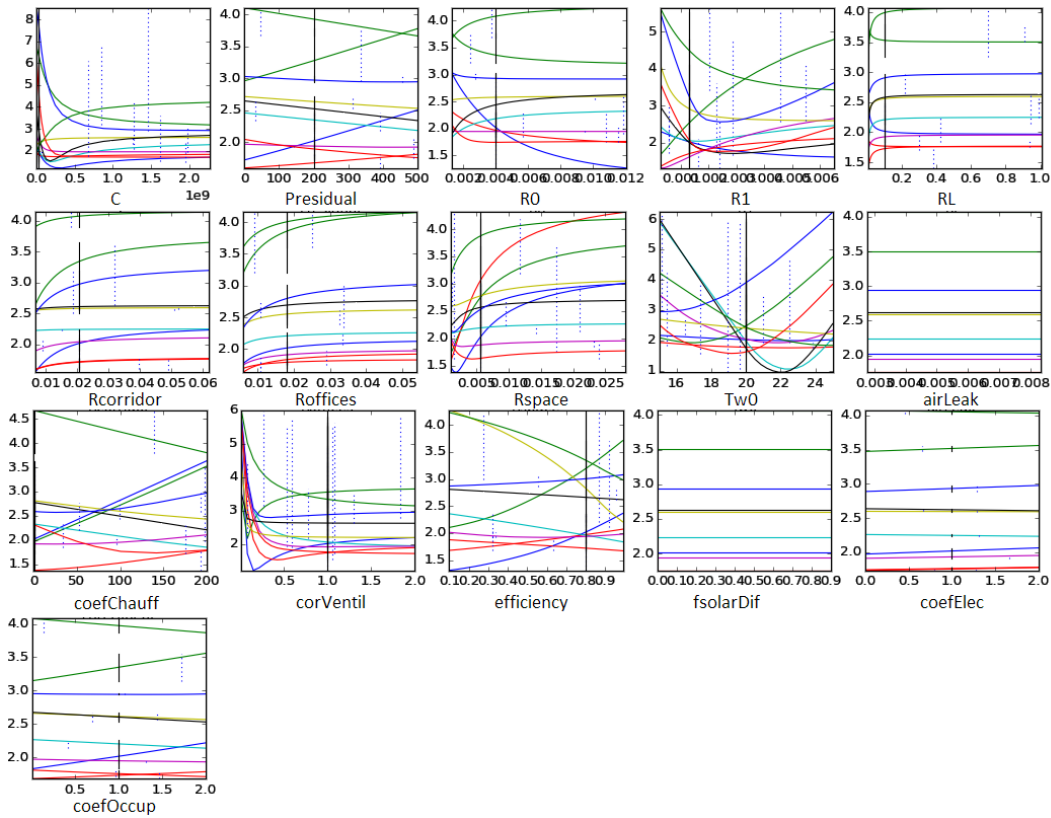


FIGURE VI.1 – Exemple d'évolution du critère à minimiser en fonction des variations de paramètres pour tous les autres paramètres fixés

tion de la valeur des autres paramètres, l'évolution du critère est différente et la monotonie peut changer. Tandis que certains paramètres admettent régulièrement une valeur qui minimise le critère (paramètres R_1 , T_{w0}), d'autres paramètres ne sont pas ajustables (paramètres $R_{residual}$, air_{Leak} , les coefficients de flux solaire...).

A chaque début d'itération, l'algorithme va observer la monotonie du critère par rapport à chaque paramètre et juger de l'ajustabilité de celui-ci.

La figure VI.2 montre le type de courbe que l'on peut obtenir et la conclusion que l'on peut tirer sur l'ajustabilité des paramètres.

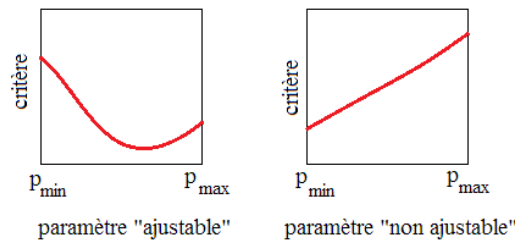


FIGURE VI.2 – Ajustabilité des paramètres en fonction de la courbe des critères

A l'issue de la phase d'analyse, les paramètres ajustables seront estimés lors de l'optimisation tandis que les paramètres non ajustables seront fixés à leur valeur courante : déduite d'une connaissance physique ou estimée à un pas précédent où le paramètre était ajustable. Ceci permet de ne pas repartir de zéro lorsqu'une valeur ajustable a été trouvée une fois.

VI.1.2-ii Optimisation : Estimation paramétrique des paramètres ajustables

Lors de la phase d'optimisation, les paramètres jugés ajustables lors de l'analyse sont estimés. On obtient un vecteur de paramètres composé des paramètres ajustables estimés et des paramètres non ajustables à leur valeur courante.

Si certains paramètres estimés sont arrivés hors du domaine acceptable alors le modèle est rejeté et ils sont ramenés à leur valeur courante (valeur a priori). Une nouvelle itération peut commencer. Ce recalage itératif des paramètres permet de guider l'algorithme dans l'espace de recherche.

La méta-optimisation est calibrée pour réaliser un certain nombre d'itérations. Contrairement à la méthode décrite au chapitre précédent, ici, nous ne cherchons pas à fixer les paramètres définitivement. De plus, la manière de parcourir l'espace de recherche est différente. En effet, tandis qu'à chaque itération, l'initialisation des paramètres se faisaient par tirage aléatoire dans l'espace de recherche, ici, l'initialisation à chaque début d'itération est liée au point d'arrivée précédent, où les valeurs cohérentes sont conservées et les valeurs aberrantes (hors du domaine acceptable) sont ramenées à leur valeur courante précédente.

Enfin, si l'algorithme retombe sur un ensemble de valeurs de paramètres qui a déjà été trouvé dans une étape d'optimisation, alors une diversification est réalisée afin de partir d'un point de départ aléatoire différent pour parcourir au mieux l'espace de recherche.

VI.1.2-iii Choix du meilleur modèle

A la fin de la méta-optimisation, parmi les modèles acceptés (dont aucune valeur de paramètre identifiée n'est hors du domaine acceptable), le modèle retenu est celui dont l'erreur moyenne est la plus faible.

L'évaluation du modèle au point de départ pour les valeurs nominales des paramètres donne d'office un modèle acceptable. Si les modèles acceptables trouvés pendant la méta-optimisation ne présentent pas une erreur moyenne plus faible que ce modèle, alors le modèle initial sera conservé.

VI.1.3 Algorithme

Voici comment procède l'algorithme qui pilote l'estimation paramétrique :

```

Collecter fonctions utilisateurs simulate() et objectif()
Evaluer objectif pour jeu de paramètres initial
Ajouter jeu de paramètres initial à l'historique des jeux de paramètres

meilleur jeu de paramètres = jeu de paramètres initial

Tant que nombre max de pas méta-optimisations <  $Nombre_{max}$  ET des paramètres
peuvent être ajustés faire

```

Déterminer les valeurs des paramètres qui doivent être ajustés

Si certains paramètres peuvent être ajustés alors

Ajuster valeur des paramètres avec algorithme d'optimisation

Si nouveau jeu de paramètres est dans l'historique alors

Diversifier valeurs de paramètres non fixés par tirage
aléatoire dans les bornes de recherche

Ajuster valeurs de paramètres avec algorithme d'optimisation
en utilisant valeurs diversifiées comme point de départ

fin si

Ajouter nouveau jeu de paramètres à l'historique

fin si

Si nouveau jeu de paramètres acceptable ET nouvelles valeurs donnent
meilleur objectif que meilleur jeu de paramètres alors

meilleur jeu de paramètres = jeu de paramètres courant

fin si

Si certaines valeurs de paramètres sont en butée du domaine de validité
alors

remettre valeurs de paramètres aux valeurs courantes

fin si

fin tant que

VI.1.4 Pré-requis

Contrairement à la précédente méthode, cet algorithme ne nécessite pas d'études préalables des paramètres. En effet, l'algorithme s'utilise en temps réel. A chaque estimation paramétrique, 20 analyses et optimisations successives sont réalisées. Il s'agit d'identifier les paramètres du modèle à un moment précis en vu de faire de la gestion anticipative.

Cependant, de même que précédemment, avant de mettre en place ce genre d'algorithme, il faut posséder un modèle du bâtiment, et avoir défini des valeurs a priori des paramètres ainsi que les bornes qui délimitent l'espace de recherche.

Cette méthode a été testée sur le modèle des salles PREDIS MHI et validée sur un autre modèle de bâtiment.

VI.2 Mise en place de la méthode sur des jeux de données PREDIS MHI

Comme pour la méthode du chapitre précédent, celle-ci sera dans un premier temps testée sur des jeux de données issus de la simulation via le logiciel COMFIE Pléiade puis ensuite sur deux jeux expérimentaux. Il s'agira de tester plusieurs méthodes d'estimation paramétrique et de les comparer en terme de rapidité d'estimation et de précision en prédiction.

VI.2.1 Simulation sur un an de jeux de données simulés issus du logiciel COMFIE

Nous allons ainsi vérifier que la méta-optimisation permet une recalibration robuste des jeux de données dans la cadre de la gestion anticipative. Elle va donc être comparée avec une méthode de calibration classique (estimation paramétrique donnant l'erreur la plus faible parmi n) et à un algorithme génétique. Plusieurs jeux de données ont été extrait de l'année complète pour être identifiés. Les bornes limitant l'espace de recherche seront les mêmes pour les deux types de calibration (classique et méta optimisation). Elles sont calculées analytiquement en fonction des caractéristiques du bâtiment et des incertitudes que nous avons sur ces caractéristiques (épaisseur du mur, coefficient de conductivité thermique, etc.). L'algorithme génétique utilise les mêmes bornes de recherche.

Dans un premier temps, les erreurs moyennes obtenues vont être comparées pour plusieurs jeux de données. Ensuite, les résultats seront validés en simulant les modèles obtenus en validation sur des scénarios différents.

VI.2.1-i Comparaison des résultats d'identification

Nous disposons de 12 jeux de données représentant les mois de l'année et issus d'une simulation du logiciel COMFIE Pléiade. Ces 12 jeux de données ont été utilisés pour identifier les paramètres du modèle avec les méthodes suivantes :

- **Calibration classique** : Parmi plusieurs optimisations (ici, 10), le modèle donnant l'erreur la plus faible est conservé. C'est la fonction Matlab *fmincon* qui est utilisée avec l'algorithme 'active-set' qui réalisent des "bonds" de moyennes échelles, ce qui permet d'augmenter la vitesse de convergence.
- **Algorithme génétique** : Utilisation de la notion de sélection naturelle appliquée à une population de solutions potentielles. La solution est approchée par "bonds" successifs comme dans une procédure de séparation et évolution.
- **Méta-optimisation** : Recalage itératif des paramètres afin d'obtenir un modèle ayant des valeurs de paramètres strictement comprises entre les bornes de l'espace de recherche. Cette méthode réalise 20 itérations qui sont chacune composées d'une phase d'analyse et d'une phase d'optimisation.

Le tableau VI.1 rend compte des valeurs d'erreurs moyennes obtenues pour les deux méthodes. Il est flagrant que la méthode d'identification utilisant une calibration classique donne une bien meilleure précision. L'algorithme génétique donne lui aussi de très bons résultats en terme d'erreur moyenne.

| | Méthode de calibration classique | Algorithme génétique | Méta-optimisation |
|----------------|----------------------------------|----------------------|-------------------|
| Janvier | 0,29 | 0,34 | 0,52 |
| Février | 0,26 | 0,56 | 0,32 |
| Mars | 0,19 | 0,38 | 0,48 |
| Avril | 0,18 | 0,42 | 0,45 |
| Mai | 0,18 | 0,13 | 1,19 |
| Juin | 0,20 | 0,12 | 0,65 |
| Juillet | 0,15 | 0,28 | 0,39 |
| Août | 0,13 | 0,22 | 0,16 |
| Septembre | 0,24 | 0,32 | 0,73 |
| Octobre | 0,18 | 0,39 | 0,26 |
| Novembre | 0,18 | 0,17 | 0,56 |
| Décembre | 0,14 | 0,12 | 0,61 |
| Moyenne | 0,19 | 0,29 | 0,53 |

TABLEAU VI.1 – Résultats d'erreurs moyennes obtenues en estimation paramétrique pour plusieurs jeux de données obtenus par simulation via le logiciel COMFIE Pléiade

C'est assez logique puisque ces méthodes possèdent plus de degré de liberté par rapport à la méta-optimisation qui contraint les valeurs de certains paramètres qui s'éloignent trop des valeurs nominales pendant l'identification.

Cependant, les valeurs des erreurs moyennes obtenues en méta-optimisation restent correctes (inférieures au degré d'erreur dans la plupart des cas). De plus, le temps de calcul du modèle par la méta-optimisation est beaucoup plus rapide, comme le montre le tableau VI.2. La méta-optimisation est donc en moyenne 6 fois plus rapide à converger que la calibration classique et environ 45 fois plus rapide qu'un algorithme génétique !

| | Calibration classique (10 itérations) | Algorithme génétique | Méta-optimisation (20 itérations) |
|---------------------------|--|----------------------|--------------------------------------|
| temps de simulation moyen | 2 140 secondes | 14 200 secondes | 326 secondes |

TABLEAU VI.2 – Temps de simulation moyen par méthode d'estimation paramétrique

Dans le cadre de la gestion anticipative, c'est la précision en prédiction qui importe puisque nous souhaitons anticiper la température intérieure le plus précisément possible afin de mettre en place des scénarios de gestion optimaux. La méta-optimisation, en guidant l'optimisation dans un domaine physique, garantit un modèle extrapolable à d'autres situations, comme nous allons le vérifier dans la section suivante.

VI.2.1-ii Comparaison des résultats en prédiction : vérification de l'extrapolabilité du modèle

Il s'agit ici de vérifier la capacité d'extrapolation du modèle lorsqu'il est identifié avec la méta-optimisation. Nous avons extrait quatre jeux de données d'une durée à peu près égale (5 ou 6 jours) dans l'année de simulation issue du logiciel COMFIE Pléiade. Ces jeux de données ont été choisis car leurs scénarios sont différents en termes de chauffage et de ventilation.

Les quatre jeux de données ont été extraits sur une période hivernale (qui fait normalement intervenir du chauffage). Certains des jeux de données ont été extraits de périodes de vacances scolaires, d'où l'absence de chauffage et de ventilation pour ces jeux de données. La figure VI.3 montre les courbes de ventilation (débit d'air) et de chauffage (puissance) de ces quatre jeux de données.

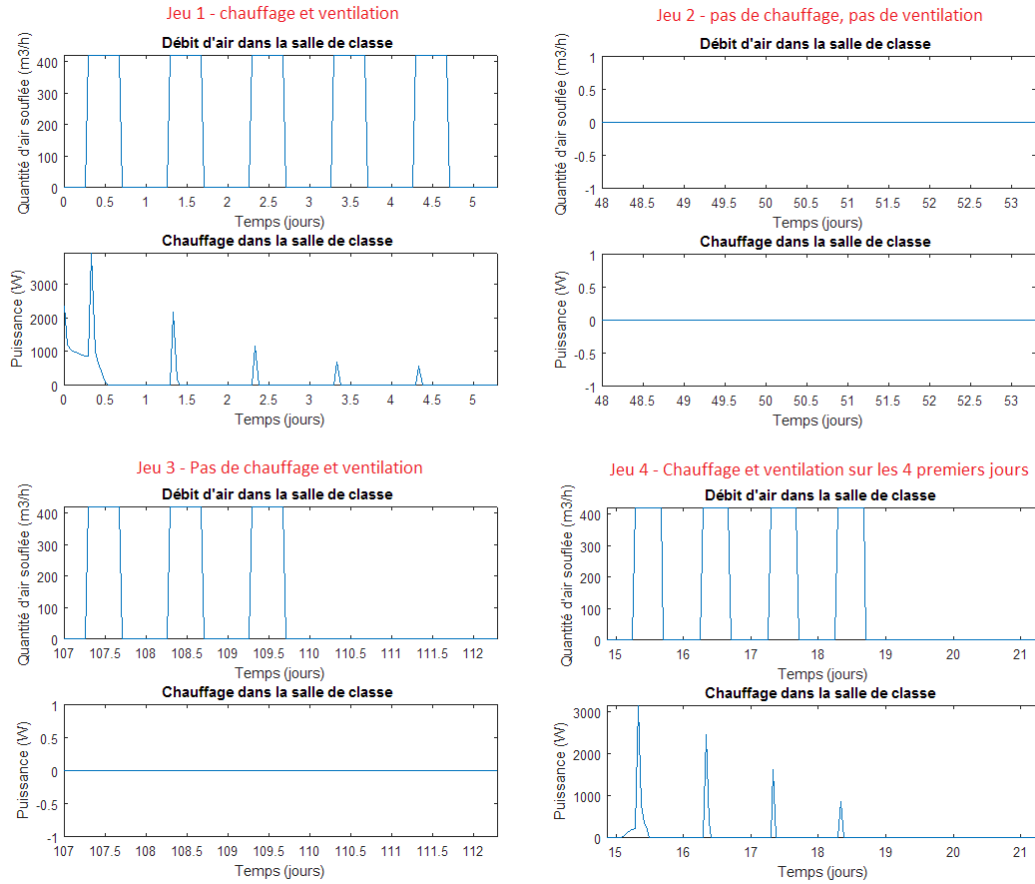


FIGURE VI.3 – Observation du débit d'air et de la puissance de chauffage pour les quatre jeux de données utilisés pour la validation

| Jeu | Chauffage | Ventilation |
|-------|--------------|--------------|
| Jeu 1 | oui | oui |
| Jeu 2 | non | non |
| Jeu 3 | non | intermittent |
| Jeu 4 | intermittent | intermittent |

TABLEAU VI.3 – Description des jeux de données utilisés pour les tests d'extrapolabilité

Il s'agit ici d'identifier les paramètres du modèle à l'aide d'un jeu de données ayant un scénario spécifique et de simuler ce modèle avec un autre jeu de données ayant un scénario différent.

Nous avons identifié le modèle d'après le jeu de données 2, qui n'a fait intervenir ni chauffage ni de ventilation. Les erreurs moyennes obtenues en identification sont visibles dans le tableau VI.4.

| | Calibration classique | Algorithme génétique | Méta-optimisation |
|----------------|-----------------------|----------------------|-------------------|
| Erreur moyenne | 0,09°C | 0,06°C | 0,21°C |

TABLEAU VI.4 – Erreurs moyennes obtenues en identification selon différentes méthodes d'estimation pour le jeu 2 (scénario sans chauffage ni ventilation)

Tous les résultats obtenus sont très bons mais les erreurs d'estimations de la calibration classique et de l'algorithme génétique sont nettement meilleurs (de 2 à 3 fois). Voyons les résultats que nous obtenons en validation avec les trois autres jeux de données.

| | Calibration classique | | Algorithme génétique | | Méta-optimisation | |
|-------|-----------------------|----------------|----------------------|----------------|-------------------|----------------|
| | Erreur moyenne | Erreur maximum | Erreur moyenne | Erreur maximum | Erreur moyenne | Erreur maximum |
| jeu 1 | 2,59 | 7,22 | 2,03 | 15,09 | 0,90 | 2,45 |
| jeu 3 | 2,79 | 6,50 | 0,84 | 4,16 | 1,12 | 3,24 |
| jeu 4 | 2,30 | 5,97 | 0,79 | 10,32 | 0,90 | 2,10 |

TABLEAU VI.5 – Résultat d'un modèle en validation pour différents scénarios de chauffage et de ventilation dont les paramètres ont été identifiés d'après un jeu sans chauffage, ni ventilation

Les erreurs moyennes et maximales obtenues en validation suite à une estimation paramétrique par une calibration classique sont assez médiocres, avec une erreur de 2,5°C en moyenne. De plus, l'erreur atteint des valeurs relativement élevées (entre 5 et 7°C). Concernant les résultats obtenus pour l'algorithme génétique, les erreurs moyennes sont relativement bonnes, mais les erreurs maximums atteintes sont très mauvaises. Les résultats obtenus par la méta-optimisation donnent quant à eux une erreur moyenne qui est largement acceptable dans le cadre d'une gestion anticipative. De plus, les valeurs d'erreur maximales atteintes sont beaucoup plus faibles que pour les autres méthodes d'estimation. Ceci démontre que lorsque le modèle est identifié avec la méta-optimisation, il possède une plus grande capacité d'extrapolation que lorsqu'il est estimé avec d'autres méthodes. Les contraintes exercées sur les bornes physiques du domaine permettent de garantir un modèle capable de représenter des phénomènes non présents dans le jeu d'identification.

Testons maintenant la méthode sur des jeux de données expérimentaux.

VI.2.2 Validation de la méta-optimisation sur les jeux de données expérimentaux des salles PREDIS MHI

Nous avons une fois de plus utilisé les huit jeux expérimentaux issus des capteurs des salles PREDIS MHI. Ces huit jeux de données ont été identifiés avec la méthode de méta-optimisation. Ici, les coefficients de flux solaires ont été fixés car l'expérience a montré qu'ils sortaient toujours du domaine acceptable. Le domaine acceptable est celui qui a été délimité par l'expert. Cette limite est subjective, bien que basée sur une connaissance physique et peut également être remise en question lors de l'étude du modèle.

VI.2.2-i Observation des résultats

La méthode de méta-optimisation oriente l'estimation paramétrique dans l'espace de recherche afin d'éviter les zones physiquement inacceptables. La figure VI.4 rend compte des résultats obtenus pour un jeu de données de janvier 2015. La méta-optimisation a été réalisée en 20 itérations.

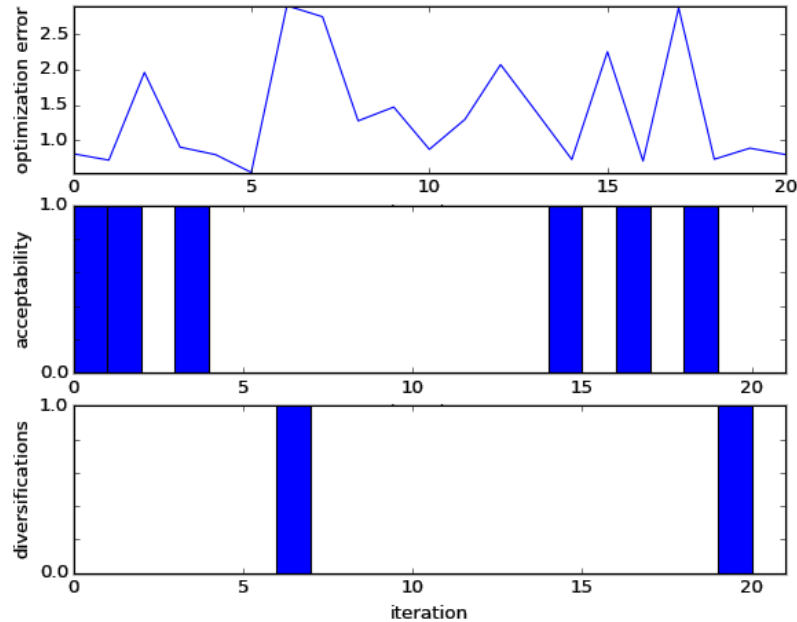


FIGURE VI.4 – Résultat de la méta-optimisation pour un jeu de janvier 2015

Le graphique du haut montre l'évolution de l'erreur moyenne en fonction de l'itération. Le graphique du milieu rend compte de la validité (vaut 1) ou non (vaut 0) du modèle à chaque itération. Un modèle est valide, ou acceptable, lorsque les valeurs de tous les paramètres (identifiés ou fixés) sont strictement incluses dans le domaine de validité. Le dernier graphique représente les itérations où une diversification a été réalisée.

On remarque sur cette figure que sur 20 itérations, 6 ont mené à un modèle acceptable. Parmi ces 6 modèles, celui conservé pour la gestion anticipative est celui qui a l'erreur moyenne la plus faible. Toutefois, certains modèles parmi les modèles rejetés présentaient une erreur moyenne plus basse que celles des modèles acceptables (modèle de l'itération 5 par exemple). C'est plutôt logique puisqu'en contraignant l'optimisation, on limite les degrés de liberté.

On remarque également que 2 diversifications ont été réalisées.

Nous allons maintenant comparer l'erreur obtenue à l'issue de la première itération de la méthode à celle obtenue à l'issue de la méta-optimisation.

VI.2.2.i - a) Erreur moyenne en identification

Chaque méta-optimisation commence par une évaluation du modèle avec les valeurs initiales des paramètres, préalablement calculées de manière analytique. Ce modèle est d'office acceptable puisque l'ensemble des valeurs des paramètres est strictement inclu dans le domaine valide. Une erreur moyenne est obtenue pour ce modèle. A l'issue de la méta-

optimisation, le modèle retenu correspond au modèle estimé donnant la meilleure erreur parmi les modèles non rejetés, c'est-à-dire ne présentant pas de paramètres dont la valeur estimée est en butée des bornes du domaine. Le graphique de la figure VI.5 permet de comparer les erreurs moyennes initiales et celle obtenue à l'issue de la méta-optimisation.

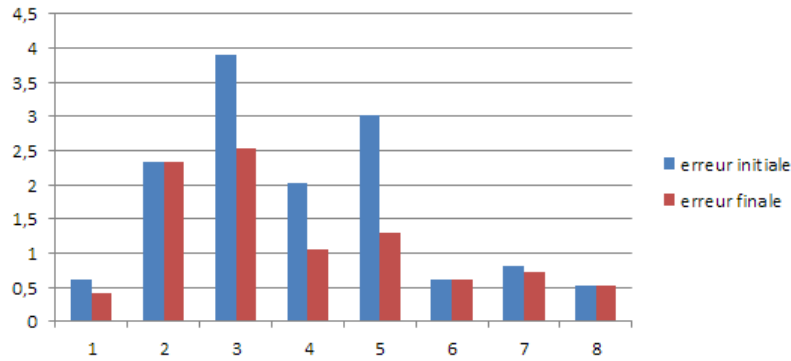


FIGURE VI.5 – Observation des erreurs obtenues pour le modèle initial et à l'issue de la méta-optimisation pour les huit jeux de données

L'erreur est toujours améliorée ou égale à l'erreur initiale (et cela ne peut être autrement). Lorsque l'erreur finale est égale à l'erreur initiale, cela signifie qu'aucun modèle acceptable n'a été trouvé. Ici, trois des huit jeux de données seulement conservent la même erreur à l'issue de la méta-optimisation. Ainsi, tout en garantissant des valeurs de paramètres ayant une cohérence physique, on améliore la précision de l'estimation.

Lorsque l'algorithme de méta-optimisation trouve un modèle qui minimise l'erreur tout en ayant des valeurs estimées de paramètres strictement incluses dans les intervalles de recherche, il est intéressant de regarder quels sont les paramètres qu'il juge ajustables ou non, et si ceux-ci sont les mêmes pour des jeux de données différents.

VI.2.2.i - b) Cohérence des résultats

La méta-optimisation renvoie le meilleur modèle qui minimise l'écart entre mesure et simulation tout en ayant des valeurs de paramètres estimées strictement incluses dans le domaine acceptable. Ce modèle estimé contient des paramètres ajustables et des paramètres non ajustables. Ceci dépend des valeurs de chaque paramètres et donc du point de l'espace de recherche.

Si nous comparons les résultats obtenus pour chacun des huit jeux de données dans le tableau VI.6, nous remarquons que les paramètres ajustables sont souvent les mêmes. Parmi les paramètres qui sont toujours ajustables quels que soient les jeux de données utilisés, on retrouve : R_{space} , $R_{corridor}$, R_L , R_0 , R_1 , C , T_{w0} et $coef_{Ventil}$. En revanche, air_{Leak} et les coefficients de flux solaires sont des paramètres qui ne sont jamais considérés comme ajustables.

Ce genre de résultats montre une certaine cohérence quant à l'ajustabilité des paramètres. En effet, cela démontre que certains paramètres sont facilement et globalement identifiables tandis que d'autres sont globalement non identifiables. Parmi ces paramètres difficilement identifiables (peu souvent ajustables), on retrouve air_{Leak} , $R_{offices}$ et les coefficients de flux

| | Février 2014 | Avril 2014 | Juin 2014 | Juillet 2014 | Août 2014 | Décembre 2014 | Janvier 2015 | Février 2015 |
|---------------------|-----------------|---------------|--------------|-----------------|--------------|------------------|-----------------|-----------------|
| R_{space} | x | x | x | x | x | x | x | x |
| $R_{offices}$ | x | | | | | | x | |
| $R_{corridor}$ | x | x | x | x | x | x | x | x |
| R_L | x | x | x | x | x | x | x | x |
| R_0 | x | x | x | x | x | x | x | x |
| R_1 | x | x | x | x | x | x | x | x |
| C | x | x | x | x | x | x | x | x |
| $efficiency$ | x | x | x | x | x | | x | |
| T_{w0} | x | x | x | x | x | x | x | x |
| $P_{residual}$ | x | x | | x | x | x | x | x |
| air_{Leak} | | | | | | | | |
| $coef_{Occup}$ | x | x | | x | x | | x | x |
| $coef_{Elec}$ | x | x | | | | | x | x |
| $coef_{Chauff}$ | x | x | | x | x | x | x | x |
| $coef_{Ventil}$ | x | x | x | x | x | x | x | x |
| f_{solar}^{East} | | | | | | | | |
| f_{solar}^{South} | | | | | | | | |
| f_{solar}^{West} | | | | | | | | |
| f_{solar}^{Diff} | | | | | | | | |

TABLEAU VI.6 – Paramètres ajustables "x" et non ajustables à l'issue de la méta-optimisation

solaire. Le coefficient de fuite d'air air_{Leak} peut en effet être complexe à identifier avec une valeur fixe car le débit de ventilation évolue au cours du temps. La résistance thermique $R_{offices}$ quant à elle, est complexe à identifier car le mur qui sépare la salle de classe de l'espace bureau est très conducteur et les températures des deux pièces sont très proches. Concernant les flux solaires, ceux-ci sont interpolés en fonction de données géométriques et de la date, mais nous manquons de données météorologiques pour les modéliser précisément. Les coefficients multiplicateurs de flux solaires sont donc complexes à identifier de manière constante sur un jeu de données de plusieurs jours où la météo peut changer.

On retrouve des similitudes quant au taux d'ajustabilité par paramètre par rapport à la figure VI.1. En effet, le paramètre T_{w0} (qui représente une condition initiale) admettait un minimum pour la courbe du critère dans 75% des cas. Ici, la méta-optimisation a montré que pour chaque modèle retenu, le paramètre T_{w0} était ajustable.

Néanmoins, ce résultat montre la nécessité de relancer la méta-optimisation lorsque le jeu de données s'enrichit car les paramètres ajustables ne sont pas strictement identiques d'un jeu à l'autre.

VI.2.2-ii Comparaison des résultats d'identification

Comparons les résultats obtenus avec ceux d'une calibration classique et d'un algorithme génétique. Le tableau VI.7 montre les erreurs moyennes que nous obtenons pour les trois méthodes.

Les précisions sur le jeu d'apprentissage de la calibration classique et de l'algorithme génétique restent en moyenne meilleures que la méta-optimisation, comme c'était le cas pour les jeux de données COMFIE. L'écart est néanmoins moins marqué qu'avant.

| | Erreur moyenne Calibration classique | Erreur moyenne Algorithme génétique | Erreur moyenne Méta-optimisation |
|---------------------|---|--|-------------------------------------|
| 1-23 juin 2014 | 1, 12 | 0, 95 | 0, 87 |
| 1-26 janvier 2015 | 0, 58 | 0, 54 | 0, 72 |
| 5-18 février 2014 | 0, 91 | 0, 71 | 0, 42 |
| 7-15 juillet 2014 | 1, 04 | 0, 97 | 1, 05 |
| 7-24 avril 2014 | 0, 76 | 0, 64 | 2, 04 |
| 15-30 décembre 2014 | 0, 51 | 0, 45 | 0, 62 |
| 18-26 août 2014 | 0, 74 | 0, 51 | 0, 61 |
| 19-30 février 2015 | 0, 44 | 1, 03 | 0, 53 |
| moyenne | 0,76 | 0,72 | 0,86 |

TABLEAU VI.7 – Comparaison des erreurs moyennes obtenues pour diverses méthodes d'estimation paramétrique

Ici encore, le temps moyen de simulation est plus faible lorsqu'on utilise la méta-optimisation. En effet, en moyenne, le temps de simulation pour la méta-optimisation est de 124 secondes tandis qu'il faut 1 017 secondes pour la calibration classique, soit quasiment un rapport 10. Quant à l'algorithme génétique, il met en moyenne un peu moins de 5 heures (18 000 secondes) à converger, soit presque un rapport de 150 par rapport à la méta-optimisation ! Le temps de simulation dépend bien entendu de la longueur des jeux de données utilisés.

VI.2.2-iii Comparaison des résultats en validation

De même que précédemment, nous avons comparé les résultats en validation pour les trois méthodes de calibration. Les résultats en estimation paramétriques de certains jeux d'hiver ont été simulés sur d'autres jeux de données d'hiver. Les résultats obtenus sont présentés dans le tableau VI.8.

| | | Validation pour la calibration classique | | | | | |
|----------------|----------------|--|----------------|----------------|----------------|----------------|----------------|
| | | 1-26 janvier | | 15-30 décembre | | 19-30 février | |
| | | erreur moyenne | erreur maximum | erreur moyenne | erreur maximum | erreur moyenne | erreur maximum |
| Identification | 1-26 janvier | x | x | 0,67 | 4,73 | 0,77 | 5,32 |
| | 15-30 décembre | 1,22 | 4,51 | x | x | 0,87 | 3,94 |
| | 19-30 février | 1,34 | 4,79 | 0,61 | 2,32 | x | x |
| | | Validation pour l'algorithme génétique | | | | | |
| | | 1-26 janvier | | 15-30 décembre | | 19-30 février | |
| | | erreur moyenne | erreur maximum | erreur moyenne | erreur maximum | erreur moyenne | erreur maximum |
| Identification | 1-26 janvier | x | x | 0,73 | 4,59 | 0,84 | 5,18 |
| | 15-30 décembre | 0,94 | 4,13 | x | x | 0,75 | 3,35 |
| | 19-30 février | 1,21 | 4,87 | 0,55 | 2,11 | x | x |
| | | Validation pour la méta-optimisation | | | | | |
| | | 1-26 janvier | | 15-30 décembre | | 19-30 février | |
| | | erreur moyenne | erreur maximum | erreur moyenne | erreur maximum | erreur moyenne | erreur maximum |
| Identification | 1-26 janvier | x | x | 0,86 | 2,43 | 1,11 | 5,36 |
| | 15-30 décembre | 0,78 | 3,2 | x | x | 0,73 | 2,65 |
| | 19-30 février | 0,78 | 3,18 | 0,62 | 2,49 | x | x |

TABLEAU VI.8

Le tableau VI.9 présente les valeurs moyennes du tableau précédent. Ainsi, nous remarquons que l'erreur moyenne en validation de la méta-optimisation est meilleure que la méthode de calibration classique, et légèrement meilleure que celle de l'algorithme génétique. Cependant, l'erreur maximum atteinte est quant à elle vraiment plus faible dans le cas de la méta-optimisation. De plus, cette méthode d'estimation paramétrique est bien plus rapide à calibrer.

| | Erreur moyenne | Erreur maximum |
|----------------------------------|----------------|----------------|
| Méthode de calibration classique | 0,91 | 4,27 |
| Algorithme génétique | 0,84 | 4,04 |
| Méta-optimisation | 0,81 | 3,22 |

TABLEAU VI.9 – Valeurs moyennes des résultats d'erreur de validation pour les trois méthodes de calibrations

VI.3 Validation de la méthode avec un autre modèle de bâtiment

Il s'agit du modèle du bureau du laboratoire G-SCOP décrit dans le chapitre précédent, partie V.3.1.

La méta-optimisation a été testée et comparée à d'autres techniques d'estimations paramétriques (ARX à l'ordre 0 et à l'ordre 1) afin de vérifier que la méthode propose une recalibration robuste dans le cadre d'une gestion anticipative.

Pour les trois méthodes, une estimation paramétrique a été réalisée des données enregistrées au cours du mois de Mai et les modèles obtenus ont été simulés sur des données enregistrées au mois de Juin.

Le tableau VI.10 montre les erreurs que nous obtenons en validation pour les trois méthodes.

| | ARX d'ordre 1 | ARX d'ordre 0 | Méta-optimisation |
|----------------|---------------|---------------|-------------------|
| Erreur moyenne | 1,04 | 1,03 | 0,45 |

TABLEAU VI.10

De plus, si on regarde les courbes de températures obtenues, on remarque que lorsqu'on utilise des modèles non physiques, la courbe de simulation finit par s'éloigner de plus en plus de la mesure, comme en témoignent les figures VI.6, VI.7 et VI.8.

VI.4 Simulation de la méthode en scénario de gestion

Comme pour le chapitre précédent, cette méthode va être utilisée en simulation sur un scénario de gestion anticipative. Il s'agit ici de valider une méthode d'estimation paramétrique et non un modèle comme c'était le cas précédemment.

Divers types de scénarios ont été testés et comparés :

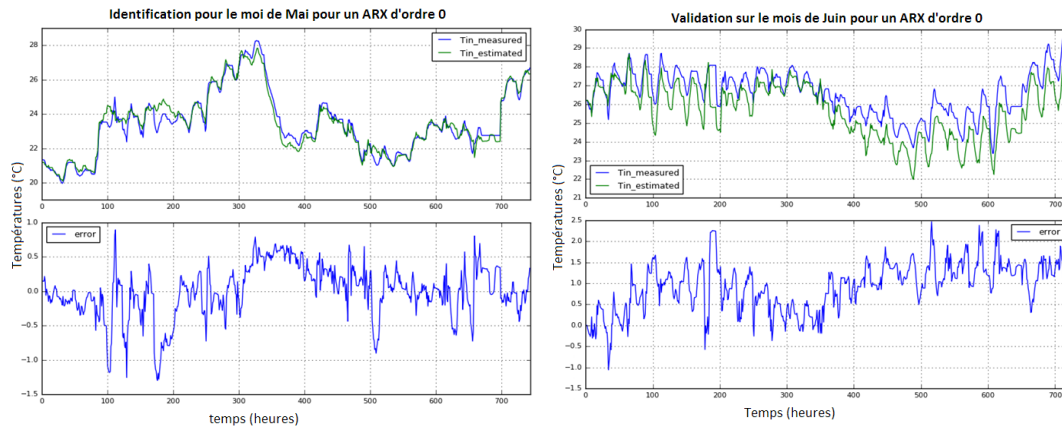


FIGURE VI.6 – Identification et validation du modèle ARX d'ordre 0 sur les mois de Mai et Juin

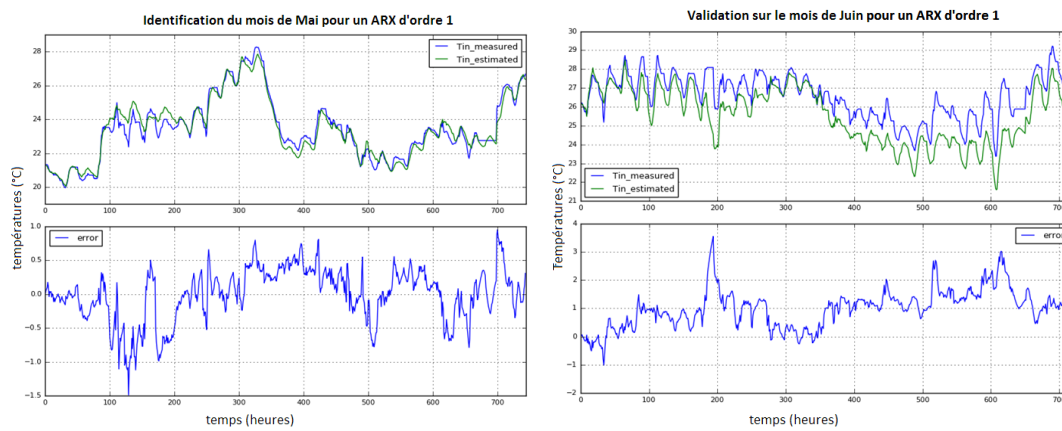


FIGURE VI.7 – Identification et validation du modèle ARX d'ordre 1 sur les mois de Mai et Juin

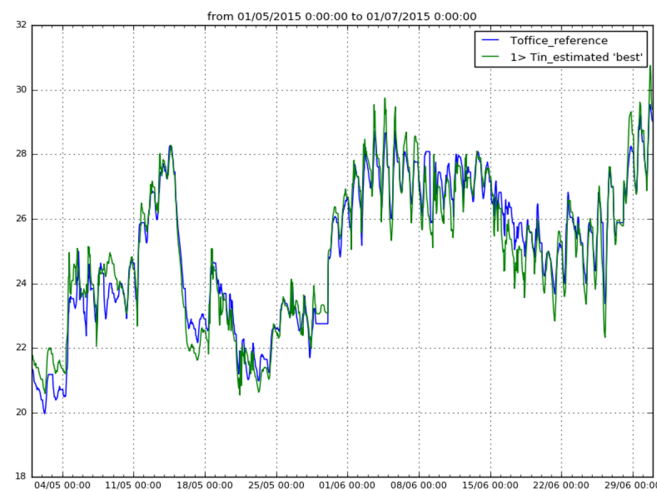


FIGURE VI.8 – Identification et validation pour le modèle physique avec la méthode de méta-optimisation

- **Scénario physique standard** : le modèle physique initial (avec l'ensemble des paramètres à identifier) est régulièrement recalibré pour la gestion anticipative à l'aide d'une méthode d'*estimation paramétrique standard* (meilleur résultat parmi 10) ;

- **Scénario physique avancé** : le modèle physique initial (avec l'ensemble des paramètres à identifier) est régulièrement recalibré pour la gestion anticipative à l'aide de la *méta-optimisation* ;
- **Scénarios non physique** : plusieurs *modèles universels* ont été testés pour la gestion anticipative, plusieurs ordres ont également été testés.

Le tableau VI.11 présente les résultats obtenus pour les diverses phases en fonction des scénarios :

Scénario physique standard

| | Erreur moyenne | Erreur maximum |
|-------------------------|----------------|----------------|
| Phases d'identification | 0,44 | 6,23 |
| Phases de validation | 0,63 | 8,38 |

Scénario physique avancé

| | Erreur moyenne | Erreur maximum |
|-------------------------|----------------|----------------|
| Phases d'identification | 0,3 | 3,75 |
| Phases de validation | 0,68 | 2,23 |

Utilisation d'un modèle ARX d'ordre 1

| | Erreur moyenne | Erreur maximum |
|-------------------------|----------------|----------------|
| Phases d'identification | 8,05 | 22,87 |
| Phases de validation | 8,12 | 22,42 |

Utilisation d'un modèle ARX d'ordre 2

| | Erreur moyenne | Erreur maximum |
|-------------------------|----------------|----------------|
| Phases d'identification | 2,81 | 25,54 |
| Phases de validation | 3,2 | 27,78 |

Utilisation d'un système d'état d'ordre 1

| | Erreur moyenne | Erreur maximum |
|-------------------------|----------------|----------------|
| Phases d'identification | 3,65 | 97,66 |
| Phases de validation | 4,49 | 96,6 |

TABLEAU VI.11 – Erreurs moyennes et maximum obtenues pour différentes phases de la gestion pour plusieurs scénarios

Tandis que les modèles non physiques donnent des résultats très mauvais, les scénarios physiques donnent de bons résultats pour de la gestion anticipative. De plus, pour une erreur moyenne en validation similaires, l'erreur maximum atteinte est bien plus faible pour la méta-optimisation, ce qui garantit une robustesse pour la gestion anticipative.

VI.5 Conclusion

Le fait de ne pas accepter les valeurs aux bornes du domaine de recherche limite les degrés de liberté de l'algorithme et peut donc donner des résultats avec une erreur moindre mais permet une validation physique garante d'une plus grande capacité à extrapoler.

La méthode de méta-optimisation garantit une cohérence physique puisque les valeurs de paramètres estimées sont strictement incluses dans les bornes de recherche du domaine et écarte donc les valeurs qui ne sont pas considérées comme acceptables.

Aucune étude préalable sur les paramètres n'est nécessaire pour utiliser cette méthode.

La méta-optimisation est plus rapide à converger qu'une méthode d'estimation standard (choisir le meilleur résultat parmi 10 identifications) puisqu'à chaque itérations, les paramètres considérés comme non ajustables à l'issue de l'analyse précédente sont fixés. Ainsi, l'optimisation réalisée sur un modèle réduit converge environ six fois plus rapidement qu'une calibration classique.

VI.6 Couplages des deux méthodes

Les deux méthodes présentées précédemment ont été testées et validées sur plusieurs modèles de bâtiments. Ces deux méthodes complémentaires peuvent offrir un bon moyen de recalibrer le modèle physique d'un bâtiment de manière robuste en vue de la gestion anticipative.

Tandis que l'une des méthodes fournit un modèle réduit dont la dispersion est limitée, ce qui permet un re-calibrage rapide et fiable, l'autre méthode fournie quant à elle une méthode d'optimisation qui mène à une estimation paramétrique physiquement cohérente.

Ces deux méthodes se focalisent sur des aspects différents de l'identifiabilité des paramètres. Néanmoins, elles se complètent dans la recherche de cohérence physique du modèle. Ainsi, il peut-être intéressant de les coupler avec dans un premier temps, une étude de la dispersion paramétrique pour fixer les paramètres les plus dispersés suivie d'une méta-optimisation pour garantir que les paramètres identifiés n'arrivent pas en butée.

VI.6.1 Principe

Commençons par réduire le nombre de paramètres à identifier.

VI.6.1-i Réduction du modèle par analyse de la dispersion paramétrique

Si nous reprenons les résultats du chapitre V sur les jeux expérimentaux de PREDIS MHI, alors les paramètres fixés sont les suivants en fonction de la saison (période de chauffage ou non) :

- Pour les jeux d'été (jeux de données sans chauffage) :
 - R_1
 - R_{space}
 - $coef_{Ventil}$
 - $P_{residual}$
 - R_L
 - air_{Leak}
 - C
- Pour les jeux d'hiver (jeux de données avec chauffage) :

- R_1
- R_{space}
- R_L
- $coef_{occup}$
- $efficiency$
- $coef_{Ventil}$
- $coef_{chauff}$

Pour la méta-optimisation, comme il avait été précisé au chapitre VI, les coefficients de flux solaires sont toujours fixés car des nombreux tests ont montré qu'ils sortaient systématiquement du domaine de recherche acceptable.

Comparons maintenant les résultats d'estimation paramétrique par méta-optimisation entre ce modèle réduit et un modèle avec l'ensemble des paramètres à identifier.

VI.6.1-ii Estimation paramétrique du modèle réduit par méta-optimisation

La méta-optimisation réalise 20 itérations. A l'issue de la méta-optimisation, un certain nombre de modèles sont acceptables (tous les paramètres identifiés sont strictement inclus dans le domaine de recherche valide). Parmi ces modèles, celui qui donne l'erreur moyenne la plus faible est utilisé en gestion anticipative.

Le tableau VI.12 rend compte de l'erreur finale obtenue pour les deux méthodes (méta-optimisation seule et couplage avec une analyse de la dispersion).

| | Méta-optimisation seule | Méta-optimisation + analyse de la dispersion |
|---------------|-------------------------|--|
| Juin 2014 | 3,18 | 3,75 |
| Août 2014 | 1,05 | 0,8 |
| Juillet 2014 | 1,05 | 0,98 |
| Avril 2014 | 2,34 | 0,75 |
| Janvier 2015 | 0,72 | 0,78 |
| Décembre 2014 | 0,62 | 0,44 |
| Février 2014 | 0,43 | 0,45 |
| Février 2015 | 0,53 | 0,43 |

TABLEAU VI.12 – Comparaison de résultats en erreur moyenne pour un couplage des deux méthodes

Nous pouvons remarquer que lorsque la méta-optimisation est précédée de l'analyse de dispersion la valeur de l'erreur finale est meilleur dans 75 % des cas, et si elle n'est pas meilleure, elle est très proche (à quelques centièmes de degrés près).

Ces premiers résultats montrent que le couplage des deux méthodes est avantageux en terme de précision. De plus, le modèle ayant moins de paramètres à estimer sera plus rapide à recalibrer par la méta-optimisation.

Cette méthode ayant pour but d'être utilisée en gestion anticipative, elle a été validée en scénario de gestion à l'aide d'un an de jeux de données simulées obtenues avec le logiciel COMFIE Pléiades.

Conclusion générale

La gestion anticipative du bâtiment permet de proposer des scénarios optimaux de fonctionnement des appareils électriques afin de réduire la consommation énergétique. Cette anticipation nécessite des modèles thermiques dynamiques dont les paramètres sont régulièrement estimés en fonction des signaux disponibles à la mesure. Le choix du modèle et de sa complexité, ainsi que la richesse des signaux d'entrée du modèle font partie des facteurs qui affectent l'identifiabilité du modèle. Or, dans le cadre de la gestion anticipative, nous cherchons des modèles robustes qui sauront s'adapter à la variabilité des sollicitations mais qui seront toutefois assez légers pour être régulièrement recalibrés.

Ainsi, mes travaux de thèse se sont focalisés sur l'étude des problématiques liées à l'identification des modèles. Ces travaux ont été appliqués aux salles PREDIS MHI, composées d'une salle de classe et d'un espace bureaux, situées dans les bâtiments du G2Elab (laboratoire de génie électrique de Grenoble) et équipées de nombreux capteurs. Les salles PREDIS MHI ont été modélisées par un réseau de résistances et de capacités thermiques.

Afin de résoudre les problématiques d'identification, nous nous sommes dans un premier temps concentrés sur l'étude du modèle, en observant la dispersion paramétrique. Ceci nous a permis d'observer que certains paramètres donnaient des résultats très dispersés sur le domaine de validité à l'issue de plusieurs optimisations. En fixant les paramètres les plus dispersés à leur valeur nominale, nous avons pu constater une amélioration de la valeur de l'erreur et une réduction de sa dispersion à l'issue d'une série d'optimisations. Plusieurs jeux de données (issus de mesures et de simulation de logiciels thermiques) ont été étudiés. Dans un premier temps, une analyse de la dispersion a permis d'observer l'évolution de la dispersion paramétrique et de l'erreur en fonction du nombre de paramètres à identifier, les autres étant fixés à leurs valeurs nominales obtenues par calcul. Un compromis en termes de nombre de paramètres à identifier a pu être trouvé. En revanche, les paramètres fixés varient en fonction des saisons. Un scénario de gestion a donc été proposé et simulé à partir d'une année de jeux de données simulées issues du logiciel de simulation thermique COMFIE Pléiade. Ainsi, le modèle réduit est utilisé dans le cadre de la gestion anticipative et les paramètres non fixés sont régulièrement ré-identifiés. Le choix des paramètres à fixer est révisé tous les trois mois. Ce scénario a été comparé à un scénario standard où le modèle initial (avec l'ensemble des paramètres) est régulièrement ré-identifié. Cette expérience a montré que l'analyse de la dispersion donnait de meilleurs résultats en terme de temps de calcul et de précision en validation. Cette méthode a également été validée sur un autre modèle de bâtiment.

Dans un second temps, nous nous sommes focalisés non plus sur le modèle mais sur la méthode d'estimation paramétrique. Nous nous sommes intéressés ici à la cohérence physique du modèle. Des bornes délimitant l'espace de recherche paramétrique ont été définies. Ces bornes permettent de conserver la cohérence physique du modèle. Toutefois, afin de ne pas trop contraindre l'estimation, elles ont été choisies plus ou moins larges en fonction des incertitudes que nous avons sur le calcul des paramètres. En acceptant que l'algorithme atteigne les bornes du domaine de recherche, nous perdons en cohérence physique. C'est pourquoi nous avons défini un algorithme qui contraint les valeurs des paramètres estimés afin qu'ils n'atteignent pas les bornes. A chaque itération, à l'issue d'une optimisation pour minimiser l'écart entre mesure et simulation, les paramètres étant arrivés aux bornes de l'espace de recherche

sont ramenés à leurs valeurs nominales ou à la dernière valeurs estimées non arrivée en butée de l'espace de recherche. A l'issue d'une vingtaine d'itérations, le modèle présentant l'erreur la plus faible parmi les modèles acceptables (n'ayant aucune valeur de paramètres estimée en bornes du domaine de recherche) est retenu. Cette méthode d'estimation paramétrique est utilisée pour régulièrement recalibrer le modèle dans le cadre d'une gestion anticipative. Le fait de contraindre l'algorithme en empêchant les valeurs des paramètres de trop s'éloigner des valeurs physiques donne de moins bons résultats en identification qu'une méthode de calibration classique, mais permet une grande capacité d'adaptation du modèle en terme de prédiction avec des jeux de données ayant des sollicitations différentes. Les résultats sur une année ont montré que la méthode de méta-optimisation donnait de meilleurs résultats qu'une calibration classique ou qu'un modèle universel. Cette méthode a également été validée sur un autre modèle de bâtiment.

Ainsi, ces deux méthodes ont montré de bons résultats pour répondre aux enjeux de la gestion anticipative, c'est-à-dire permettre un recalibrage rapide et robuste du modèle du bâtiment. C'est deux méthodes, complémentaires, peuvent être couplées, mais cela affecte d'autant plus la complexité du modèle, et cet aspect reste à étudier afin de donner de meilleurs résultats.

Par la suite, il serait intéressant de valider les deux méthodes en les utilisant concrètement dans un gestionnaire énergétique de bâtiment. Des scénarios de gestion optimaux seraient déterminés en fonction des prédictions de températures obtenues par le modèle. Il s'agira de tester et de comparer plusieurs méthodes d'estimation paramétrique sur des périodes similaires (en termes de température et de météo). Ceci permettra de vérifier de l'utilité des modèles. Si l'anticipation est correcte, alors le confort sera maintenu pour un consommation énergétique minimisée.

De plus, le couplage des deux méthodes pourra être approfondi. En effet, la complexité du modèle retenu doit être liée à la richesse des signaux disponibles et le couplage des deux méthodes rend ce choix (de complexité) plus difficile. Il serait donc intéressant de l'étudier plus en détails.

ANNEXE A

Calcul des coefficients d'interface

On peut exprimer $h_{interface}$ comme ceci :

$$h_{interface} = h_{convection} + h_{rayonnement} \quad (\text{A.1})$$

- $h_{convection}$ est le coefficient thermique de convection dont la valeur dépend des situations :
 - S'il s'agit d'une paroi **intérieure**, la valeur dépend de l'orientation du flux de chaleur :
 - Horizontale (murs) : $h_{convection} = 2,5$;
 - Verticale vers le haut (plafonds) : $h_{convection} = 5$;
 - Verticale vers le bas (sols) : $h_{convection} = 0,7$.
 - S'il s'agit d'une paroi **extérieure**, alors la valeur du coefficient de convection dépend de la vitesse du vent : $h_{convection} = 4 + 4v$, avec v la vitesse du vent.
- $h_{rayonnement}$ est le coefficient thermique de rayonnement qui dépend de l'émissivité corrigée e_{corr} (celle-ci dépend de la surface), de la constante de Stefan K et de la température T en kelvin :

$$h_{rayonnement} = 4e_{corr}KT^3 \quad (\text{A.2})$$

ANNEXE B

Calculs initiaux des résistances et capacités
thermiques pour PREDIS MHI

B.1 Détails des calculs des résistances et capacité thermiques du modèle PREDIS MHI

Chaque interface est détaillée en fonction des matériaux qui la compose. Ici, nous expliciterons les détails des calculs pour l'interface "espace bureaux / salle de classe". Cette paroi est composée de mur et de vitres. Chaque type de surface est décomposée en couches.

| Matériaux | S | S_{min} | S_{max} | couche | λ | λ_{min} | λ_{max} | e | e_{min} | e_{max} |
|-----------|------|-----------|-----------|--------|-----------|-----------------|-----------------|-------|-----------|-----------|
| Vitre | 10,4 | 9,39 | 11,5 | vitre | 1,15 | 0,9 | 1,25 | 0,004 | 0,0036 | 0,0044 |
| | | | | air | 0,025 | 0,02 | 0,03 | 0,062 | 0,0558 | 0,0682 |
| | | | | vitre | 1,15 | 0,9 | 1,25 | 0,004 | 0,0036 | 0,0044 |
| Mur | 7,66 | 6,89 | 8,42 | pvc | 0,2 | 0,2 | 0,2 | 0,005 | 0,0045 | 0,0055 |
| | | | | air | 0,042 | 0,03 | 0,07 | 0,06 | 0,054 | 0,066 |
| | | | | pvc | 0,2 | 0,2 | 0,2 | 0,005 | 0,0045 | 0,0055 |

TABLEAU B.1 – Caractéristiques des matériaux et des couches qui composent l'interface espace bureau / salle de classe

avec S la surface du matériaux en m^2 et e l'épaisseur de la couche en m .

Ensuite, les coefficients d'interfaces $h_{interface}$ peuvent être calculés pour toutes les surfaces visibles :

- Pour les murs, les parois en contact avec "l'extérieur" : la salle de classe et l'espace bureau ;
- Pour les vitres :
 - interfaces en contact avec l'extérieure (salle de classe et espace bureaux) ;
 - interfaces air/vitre à l'intérieur du double vitrage ;

Le calcul des coefficients d'interface $h_{interface}$ est détaillé en annexe A.

Le tableau B.2 montre les valeurs minimum et maximum qui ont été utilisés pour les calculs dans l'exemple d'un coefficient d'interface pour une paroi pvc. Si la paroi avait été une vitre, c'est la valeur de l'émissivité corrigée qui aurait changé.

| | | Valeur minimum | Valeur nominale | Valeur maximum |
|-----------------|------------------------------------|----------------|-----------------|----------------|
| Convection | horizontale | 2,5 | 2,5 | 2,5 |
| Rayonnement | émissivité corrigée | 0,85 | 0,9 | 0,97 |
| | Cst de Stefan | $5,60e^{-8}$ | $5,60e^{-8}$ | $5,60e^{-8}$ |
| | T (degresC) | 18 | 22 | 26 |
| | T (K) | 291,5 | 295,15 | 299,15 |
| $h_{interface}$ | $h_{convection} + h_{rayonnement}$ | 7,20 | 7,68 | 8,32 |

TABLEAU B.2 – Calcul des valeurs minimum et maximum d'un coefficient d'interface pour une paroi en pvc

Le tableau B.3 donne les valeurs des coefficients d'interfaces obtenus.

Ensuite, nous avons besoin du volume (V), de la densité (d) et du C_p des matériaux afin de calculer le coefficient H/T (voir tableau B.4).

Le coefficient H/T est égale à la multiplication des trois valeurs ($H/T = VdC_p$).

| matériaux | interface | h | h_{min} | h_{max} |
|-----------|---------------------------|------|-----------|-----------|
| Vitre | espace bureaux / vitre 1 | 7,8 | 4,16 | 8,20 |
| | vitre 1 / air | 7,8 | 4,16 | 8,20 |
| | air / vitre 2 | 7,8 | 4,16 | 8,20 |
| | vitre 2 / salle de classe | 7,8 | 4,16 | 8,20 |
| Mur | espace bureaux / paroi | 7,68 | 7,20 | 8,32 |
| | paroi / salle de classe | 7,68 | 7,20 | 8,32 |

TABLEAU B.3 – Valeurs des coefficients d’interface pour les matériaux de la paroi salle de classe / espace bureaux

| | | V | V_{min} | V_{max} | d | d_{min} | d_{max} | C_p | C_p^{min} | C_p^{max} |
|-------|-------|--------|-----------|-----------|------|-----------|-----------|-------|-------------|-------------|
| Vitre | vitre | 0,0417 | 0,0338 | 0,0505 | 2530 | 2280 | 2780 | 720 | 648 | 792 |
| | air | 0,647 | 0,542 | 0,782 | 1,18 | 1,22 | 1,14 | 1004 | 903,6 | 1104,4 |
| | vitre | 0,0417 | 0,0338 | 0,0505 | 2530 | 2280 | 2780 | 720 | 648 | 792 |
| Mur | pvc | 0,0383 | 0,031 | 0,0463 | 1250 | 1190 | 1350 | 1046 | 941,4 | 1150,6 |
| | air | 0,459 | 0,372 | 0,556 | 1,18 | 1,22 | 1,14 | 1004 | 903,6 | 1104,4 |
| | pvc | 0,0383 | 0,031 | 0,0463 | 1250 | 1190 | 1350 | 1046 | 941,4 | 1150,6 |

TABLEAU B.4 – Volume, densité et coefficient C_p de la paroi salle de classe / espace bureaux

Une fois toutes ces valeurs explicitées, on obtient les valeurs des résistances thermiques de toutes les couches qu’on peut additionner (car elles sont en séries) pour obtenir la résistance thermique de la paroi.

Ainsi on obtient :

$$R_{matériaux} = \frac{\sum R_{couche}}{S_{matériaux}} \quad (B.1)$$

Soit ici :

$$\begin{aligned}
 R_{vitre} &= (R_{bureau/vitre1} + R_{vitre1} + R_{vitre1/air} + R_{air} + R_{air/vitre2} + R_{vitre2} + R_{vitre2/salledeClasse}) / S_{vitre} \\
 &= \left(\frac{1}{h_{vitre1}} + \frac{e_{vitre1}}{\lambda_{vitre1}} + \frac{1}{h_{vitre1/air}} + \frac{e_{air}}{\lambda_{air}} + \frac{1}{h_{air/vitre2}} + \frac{e_{vitre2}}{\lambda_{vitre2}} + \frac{1}{h_{vitre2}} \right) / S_{vitre} \\
 R_{mur} &= (R_{bureau/paroi} + R_{pvc} + R_{air} + R_{pvc} + R_{paroi/salledeClasse}) / S_{mur} \\
 &= \left(\frac{1}{h_{bureau}} + \frac{e_{pvc}}{\lambda_{pvc}} + \frac{e_{air}}{\lambda_{air}} + \frac{e_{pvc}}{\lambda_{pvc}} + \frac{1}{h_{salledeClasse}} \right) / S_{mur} \\
 R_{paroi} &= R_{vitre} + R_{mur}
 \end{aligned} \quad (B.2)$$

Bibliographie

- [1] G. ACHTERBOSCH AND AL., *The development of a convenient thermal dynamic building model*, Energy and Buildings, (1985).
- [2] ADEME, *Chiffre cle du batiment*, 2013.
- [3] ASHRAE, *Ashrae Handbook, fundamentals*, 1981.
- [4] —, *Thermal Environmental Conditions for Human Occupancy*, 2013.
- [5] G. BRAGER AND R. DE DEAR, *Climate, comfort, and natural ventilation : A new adaptive comfort standard for ashrae standard 55*, center for the built environment, Center for the built Environment, (2001).
- [6] H.-A. DANG, B. DELINCHANT, S. GAALLOUL, AND F. WURTZ, *Modélisation électrique d'une vmc double flux et optimisation de sa commande pour l'efficacité énergétique*, IBPSA France, (2010).
- [7] S. DAUTIN, *Reduction de modeles thermiques de bâtiments : amélioration des techniques par modelisation des sollicitations meteorologiques*, 1997.
- [8] G. DEMOMENT, *Problemes inverses en traitement du signal et de l'image*, Journal de physique IV, (2002).
- [9] F. DEQUE, *Techniques numeriques de reduction de modeles lineaires : etat de l'art*, Rapport technique, (1994).
- [10] S. DOBRE, *Analyse de sensibilité et d'identifiabilité globales. application à l'estimation de paramètres photophysiques en thérapie photodynamique*, 2010.
- [11] D. DOCHAIN AND P. VANROLLEGHEM, *Automatique des bioprocédés, chapitre identification de modèles de bioprocédés, pages 53-86*, Hermes Science Publications, (2001).
- [12] J. DUFFIE, A. AND W. BECKMAN, A., *Solar Engineering of Thermal Processes*, 2006.
- [13] H. DUONG AND I. LANDAU, *An iv based criterium for model order selection*, Automatica vol.32, (1996).
- [14] P. FANGER, *Thermal comfort : analysis and applications in environmental engineering*, Automatica, vol. 32, (1996).
- [15] G. FRAISSE, V. CHRISTELLE, O. LAFABRIE, AND G. ACHARD, *Development of a simplified and accurate building model based on electrical analogy*, Energy and Buildings, (2002).
- [16] Y. GAO AND AL., *Reduced linear state model of hollow blocks walls, validation using hot box measurements*, Energy and Buildings, (2004).
- [17] M. GOUDA, S. DANAHER, AND C. UNDERWOOD, *Building thermal model reduction using nonlinear constrained optimization*, Energy and Buildings, (2002).
- [18] J. HADAMARD, *Sur les problemes aux derivees partielles et leur signification physique*, Princeton University Bulletin, (1902).
- [19] I. HAZYUK, *Dynamical optimisation of renewable energy flux in buildings*, 2011.
- [20] G. HUDSON AND C. UNDERWOOD, *A simple building modelling procedure for matlab simulink*, IBPSA, (1999).

- [21] J. IDIER, *Approche bayésienne pour les problèmes inverses*, Traite IC2, Serie traitement du signal et de l'image, (2001).
- [22] C. LATRÉMOLIÈRE, *Mise au point d'un modèle d'éclairage naturel intégré dans le gestionnaire énergétique du bâtiment canopea de la compétition solar decathlon europe 2012*, 2011.
- [23] L. LENNART, *System Identification : Theory for the user*, 1987.
- [24] S. MARSHALL, *An approximate method for reducing the order of a linear system*, Control, (1966).
- [25] N. MENDES, G. H. OLIVEIRA, AND H. X. DE ARAÚJO, *Building thermal performance analysis by using matlab/simulink*, IBPSA, (2001).
- [26] C. MENEZO, *Contribution a la modelisation du comportement thermique des batiments par couplage de modeles reduits*, 1999.
- [27] R. MERHEB, *Fiabilité des outils de prevision du comportement des systemes thermiques complexes*, 2013.
- [28] B. MOORE, *Principal component analysis in linear systems : controllability, observability, and model reduction*, IEEE Transactions on Automatic Control, (1991).
- [29] M. D. MORRIS, *Factorial sampling plans for preliminary computational experiments*, Technometrics, vol. 33, (1991).
- [30] G. PANDRAUD AND S.-G. ISOVER, *Rapid building thermal diagnosis : presentation of the qub method*, 2014.
- [31] H. PENG AND AL., *Stability analysis of the rbf-arx model based nonlinear predictive control*, Norwegian University of Science and Technology, (2012).
- [32] D. PETIT, *Reduction de modeles de connaissance et identification de modeles d'ordre reduit*, 1991.
- [33] A. SALTELLI, K. CHAN, AND E. SCOTT, *Sensitivity Analysis*, 2008.
- [34] C. TURNER AND M. FRANKEL, *Energy performance of leed for new construction buildings*, U.S. Green Building Council, (2008).
- [35] P. VAN OVERSHEE AND B. DE MOOR, *Subspace identification for linear systems : theory implementation, applications*, Technometrics, vol. 33, (1996).
- [36] E. WALTER AND L. PRONZATO, *Identification de modèles paramétriques à partir de données expérimentales*, 1994.
- [37] S. WANG AND X. XU, *Simplified building model for transient thermal performance estimation using ga-based parameter identification*, International Journal of Thermal Sciences, (2006).
- [38] C. ZAYANE, *Identification d'un modèle de comportement thermique de bâtiment à partir de sa courbe de charge*, 2011.

Table des figures

| | | |
|-------|--|----|
| II.1 | Vue extérieure du bâtiment PREDIS MHI via le logiciel IDA-ICE | 33 |
| II.2 | Vue en coupe de PREDIS MHI via le logiciel IDA-ICE | 34 |
| II.3 | Structure des modèles de la salle de classe de PREDIS MHI. Bilan des signaux d'entrée nécessaires | 34 |
| II.4 | Relevé des températures des salles PREDIS MHI pour le mois de janvier 2015 | 35 |
| II.5 | Puissance de chauffage pour Janvier et Février 2015 | 36 |
| II.6 | Energie d'un compteur électrique pour janvier 2015 | 36 |
| II.7 | Puissance obtenue par dérivation de l'énergie électrique | 37 |
| II.8 | Puissance obtenue par moyenne glissante à partir de la dérivée de l'énergie électrique | 37 |
| II.9 | Caractérisation de la course du soleil par rapport à un point local | 39 |
| II.10 | Flux solaire direct total pour deux jeux de données été / hiver | 41 |
| II.11 | Signaux obtenus par le logiciel Comfie pour le mois de Janvier | 44 |
| II.12 | Signaux obtenus par le logiciel Comfie pour le mois de Février | 44 |
| II.13 | Identification et validation avec un jeu de données obtenu par simulation avec un système d'état d'ordre 3 | 46 |
| II.14 | Identification et validation avec un jeu de données obtenu par simulation avec un modèle ARX | 47 |
| II.15 | Signaux d'entrée expérimentaux relevés en Janvier 2015 | 48 |
| II.16 | Signaux d'entrée expérimentaux relevés en Février 2015 | 49 |
| II.17 | Prédiction sur un jeu de données expérimentales avec un système d'état d'ordre 5 | 50 |
| II.18 | Prédiction sur un jeu de données expérimentales avec un modèle ARX | 50 |
| II.19 | Pôles et zéros du modèle ARX | 51 |
| II.20 | Identification et validation d'un modèle ARX non convergeant | 52 |

| | | |
|--------|--|----|
| III.1 | Identification et validation de jeux de données issus du logiciel COMFIE avec une modélisation physique | 56 |
| III.2 | Pourcentage des insatisfaits (PPD) en fonction de la sensation thermique moyenne (PMV) | 58 |
| III.3 | Plages de température acceptée par les occupants de l'espace bureau pendant l'hiver | 58 |
| III.4 | Température de confort Juin 2014 | 60 |
| III.5 | Schéma électrique équivalent de la salle de classe PREDIS MHI | 62 |
| III.6 | Schéma de la VMC double flux | 63 |
| III.7 | Débit d'air observé pour un jeu de donnée de 37 jours (Octobre/Novembre 2012) | 64 |
| III.8 | Identification de la température du couloir pour Juin et Janvier | 67 |
| III.9 | Validation de la température du couloir pour Juin et Janvier | 68 |
| III.10 | Identification de la température de la salle de recherche pour Juin et Janvier | 68 |
| III.11 | Validation de la température de la salle de recherche pour Juin et Janvier | 69 |
| III.12 | Simulation de la température intérieure de la salle de classe avec les valeurs initiales des paramètres pour des jeux de données de Janvier, Avril et Juin | 72 |
| III.13 | Identification et validation du modèle sur des jeux de données d'été et d'hiver | 73 |
| III.14 | Comparaison de la mesure du débit d'air avec les paliers retenus pour un jeu de Janvier pour la salle de classe | 75 |
| III.15 | Identification des paramètres pour un modèle avec et sans discrétisation par paliers de la ventilation | 75 |
| III.16 | Histogramme des erreurs obtenues pour 10 identifications avec 10 points de départ différents pour un jeu d'été | 76 |
| III.17 | Histogramme des erreurs obtenues pour 10 identifications avec 10 points de départ différents pour un jeu d'hiver | 77 |
| III.18 | Dispersion paramétrique pour 10 identifications avec 10 points de départ différents pour un jeu d'hiver et un jeu d'été | 77 |
| III.19 | Meilleure identification parmi 10 et validation du modèle sur des jeux de données d'hiver et d'été | 78 |
| IV.1 | Evolution de la variable de sortie pour différentes valeurs du paramètre C | 84 |
| IV.2 | Légende des paramètres utilisés dans le tableau de sensibilité | 85 |
| IV.3 | Sensibilité locale des paramètres du modèle de PREDIS MHI | 85 |
| IV.4 | Observation de la dispersion de l'erreur pour des modèles avec tous les paramètres et des modèles réduits pour les jeux de données de Janvier et de Juin | 86 |

| | | |
|-------|--|-----|
| IV.5 | Analyse de sensibilité des paramètres du modèle pour quatre jeux de données avec la méthode de MORRIS | 87 |
| IV.6 | Meilleur résultat d'identification et observation du comportement de l'erreur à l'issue d'une série d'identifications pour le modèle initial (à 19 paramètres) et un modèle réduit (à 11 paramètres) pour un jeu de Janvier 2015 | 88 |
| IV.7 | Meilleur résultat d'identification et observation du comportement de l'erreur à l'issue d'une série d'identifications pour le modèle initial (à 19 paramètres) et un modèle réduit (à 11 paramètres) pour un jeu de Juin 2014 | 88 |
| IV.8 | Observation de la dispersion paramétrique pour des bornes larges pour un jeu de données de Juin 2014 | 89 |
| IV.9 | Signaux temporels et fréquentiels de la température extérieure et des deux salles de PREDIS MHI | 90 |
| IV.10 | Génération d'une SBPA par registre à décalage bouclé | 91 |
| IV.11 | Pilotage de la pompe à chaleur de l'espace bureau à l'aide d'une séquence binaire pseudo aléatoire | 92 |
| IV.12 | Spectre de la température de la salle de classe et de l'espace bureau pour un jeu de Juillet 2014, non excité par une séquence binaire pseudo aléatoire | 92 |
| IV.13 | Spectre de la température de la salle de classe et de l'espace bureau pour un jeu d'Août 2014, excité par une séquence binaire pseudo aléatoire | 93 |
| IV.14 | Comparaison des vecteurs d'erreurs obtenues à l'issue d'une série d'identification pour l'espace bureau et la salle de classe en Juillet et Août 2014 | 95 |
| IV.15 | Observation de l'erreur moyenne et des dispersions obtenues en fonction de la longueur du jeu d'identification | 96 |
| IV.16 | Observation de l'erreur moyenne en fonction de la longueur du jeu d'identification et du jeu de validation | 96 |
| IV.17 | Évolution de l'erreur d'identification en fonction de la largeur des bornes | 98 |
| IV.18 | Évolution des erreurs de validations en fonction de la largeur des intervalles | 99 |
| V.1 | Mise en place de la réduction du modèle par analyse de la dispersion | 109 |
| V.2 | Mise en place de la réduction du modèle par analyse de la dispersion | 110 |
| V.3 | Observation de l'évolution de la dispersion de l'erreur pour chaque saison | 112 |
| V.4 | Observation de la dispersion paramétrique moyenne pour chaque saison | 113 |
| V.5 | Légende pour la figure de la dispersion paramétrique en fonction de la saison | 114 |
| V.6 | Classement des paramètres par ordre de dispersion (du plus dispersé au moins dispersé) en fonction du moment de l'année | 114 |

| | | |
|------|--|-----|
| V.7 | Dispersion des paramètres pour les jeux de données d'hiver de PREDIS et dispersion moyenne hiver | 117 |
| V.8 | Dispersion des paramètres pour les jeux de données d'été de PREDIS et dispersion moyenne été | 118 |
| V.9 | Évolution de la dispersion de l'erreur pour les jeux d'hiver de PREDIS | 120 |
| V.10 | Évolution de l'histogramme de l'erreur en fonction du nombre de paramètres à identifier pour le jeu de décembre 2014 | 120 |
| V.11 | Évolution de la dispersion paramétrique moyenne pour les jeux d'hiver de PREDIS | 121 |
| V.12 | Comparaison des modèles à 19 et à 12 paramètres identifiés sur un jeu de février 2015 et simulés sur un jeu de décembre 2014 | 122 |
| V.13 | Comparaison des modèles à 19 et à 12 paramètres identifiés sur un jeu de décembre 2014 et simulés sur un jeu de janvier 2015 | 123 |
| V.14 | Évolution du critère de dispersion pour les jeux d'été de PREDIS | 123 |
| V.15 | Évolution de l'histogramme de l'erreur en fonction du nombre de paramètres à identifier pour le jeu d'avril 2014 | 124 |
| V.16 | Évolution de la dispersion paramétrique moyenne pour les jeux d'été de PREDIS | 124 |
| V.17 | Histogramme d'erreur pour les modèles à 19 et 12 paramètres en validation pour divers jeux d'été de PREDIS | 125 |
| V.18 | Plan en coupe vu du dessus du bureau étudié | 126 |
| V.19 | Schéma aéraulique du bureau | 127 |
| V.20 | Schéma électrique équivalent pour le bureau | 128 |
| V.21 | Évolution de la dispersion de l'erreur en fonction du nombre de paramètres à identifier | 130 |
| VI.1 | Exemple d'évolution du critère à minimiser en fonction des variations de paramètres pour tous les autres paramètres fixés | 138 |
| VI.2 | Ajustabilité des paramètres en fonction de la courbe des critères | 138 |
| VI.3 | Observation du débit d'air et de la puissance de chauffage pour les quatre jeux de données utilisés pour la validation | 143 |
| VI.4 | Résultat de la méta-optimisation pour un jeu de janvier 2015 | 145 |
| VI.5 | Observation des erreurs obtenues pour le modèle initial et à l'issue de la méta-optimisation pour les huit jeux de données | 146 |
| VI.6 | Identification et validation du modèle ARX d'ordre 0 sur les mois de Mai et Juin | 150 |
| VI.7 | Identification et validation du modèle ARX d'ordre 1 sur les mois de Mai et Juin | 150 |

| | | |
|------|---|-----|
| VI.8 | Identification et validation pour le modèle physique avec la méthode de méta-optimisation | 150 |
|------|---|-----|